

Prediksi Pergerakan Indeks Harga Saham di Indonesia dengan *Google Trends*

Muhammad Jehan Mahsan¹, Dr. Deni Saepudin, S.Si., M.Si.², Aniq Atiqi Rohmawati, S.Si. M.Si.³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹jehanmahsan@student.telkomuniversity.ac.id, ²denisaepudin@telkomuniversity.ac.id,

³aniqatiqui@telkomuniversity.ac.id

Abstrak

Indeks harga saham merupakan satuan nilai untuk mengetahui pergerakan harga suatu pasar saham atau sebagian pasar saham. Indeks harga saham berfungsi sebagai tolak ukur bagi para pemegang saham untuk mendapatkan keuntungan dalam berinvestasi. Oleh karena itu, diperlukan pemodelan prediksi indeks harga saham untuk membantu para pemegang saham dalam mengambil keputusan. Banyak faktor yang mempengaruhi naik turunnya suatu indeks harga saham. Salah satunya adalah faktor ekspektasi para penjual dan pembeli saham beserta perhatian orang pada informasi naik turunnya harga saham. *Google Search* dapat digunakan oleh para pemegang saham untuk mengetahui info tentang indeks harga saham. Grafik pencarian terhadap *keywords* indeks harga saham pada *Google Search* dapat diakses melalui layanan *Google Trends*. Data pencarian indeks harga saham di *Google Trends* merupakan cerminan perhatian orang terhadap perubahan pergerakan indeks harga saham. Pada tugas akhir ini, dilakukan pemodelan prediksi pergerakan indeks harga saham di Indonesia dengan melibatkan data *Google Trends* menggunakan metode *Support Vector Regression* (SVR). Data indeks harga saham yang digunakan adalah penutupan mingguan indeks IHSG dan LQ45. Pemodelan prediksi indeks IHSG dan LQ45 dengan melibatkan data *Google Trends* memiliki akurasi lebih baik dibandingkan pemodelan tanpa melibatkan data *Google Trends*. Namun nilai akurasi prediksi kedua model tidak signifikan berbeda.

Kata kunci : indeks harga saham, prediksi, *Google Trends*, *Support Vector Regression*, IHSG, LQ45

Abstract

Stock price index is a unit of value to determine the movement of a stock market or the most of the stock market price. Stock price index serves as a benchmark for shareholders to get profits in investing. Therefore, it is necessary to make a prediction model of stock price index to assist shareholders in making decisions. There are many factors that can affect the ups and downs of a stock price index. One of the factor is the expectation of stock sellers and buyers along with people's attention on the ups and downs of stock prices information. *Google Search* can be used by shareholders to find out info about the stock price index. Search charts of stock price index keywords on *Google Search* can be accessed on the *Google Trends* service. Stock price index search data on *Google Trends* is a reflection of people's attention to changes in stock price index movements. In this final project, modeling prediction of stock price index movements in Indonesia with involving *Google Trends* data is built using the *Support Vector Regression* (SVR) method. Stock price index data that we used is *IDX* and *LQ45*. Prediction model for *IDX* and *LQ45* index with involving *Google Trends* data has better results than prediction model without involving *Google Trends* data. However, the predicted accuracy of the two models is not significantly different.

Keywords: Stock price index, prediction, *Google Trends*, *Support Vector Regression*, IHSG, LQ45

1. Pendahuluan

Latar Belakang

Saham merupakan tanda penyertaan modal seorang atau pihak badan usaha dalam suatu perusahaan atau perseroan terbatas [1]. Sedangkan, Indeks harga saham merupakan sebuah indikator pergerakan harga saham dalam kurun waktu tertentu (time series). Banyak faktor yang dapat mempengaruhi dalam pergerakan indeks harga saham. Seperti faktor politik, faktor ekonomi, serta faktor ekspektasi para penjual dan pembeli saham. Para pemegang saham biasa menggunakan indeks harga saham sebagai tolak ukur dalam berinvestasi di pasar modal.

Memprediksi indeks harga saham memiliki kompleksitas tersendiri bagi para pemegang saham maupun investor untuk mendapatkan keuntungan yang maksimal. Dikarenakan banyaknya variabel yang terlibat serta dapat mempengaruhi harga saham baik variabel numerik ataupun kategorik. Banyak penelitian yang sudah dilakukan untuk memprediksi pergerakan indeks harga saham menggunakan metode seperti *ARIMA*, *Artificial Neural Network* (ANN), *Backpropagation Neural Networks* (BPNN), *Random Forest*, *Support Vector Machine* (SVM), dan *Support Vector Regression* (SVR) [2, 3, 4, 5, 6, 7, 8, 9, 10].

Di era digital ini, mesin pencari *Google* sering digunakan untuk mencari suatu informasi dengan menggunakan kata kunci yang berhubungan. Statistik kepopuleran pencarian kata kunci pada mesin pencari *Google* dapat dilihat melalui *Google Trends*, dimana *Google Trends* merupakan layanan yang disediakan oleh *Google* yang dapat diakses secara terbuka dan menyediakan layanan untuk melihat tingkat kepopuleran kata kunci pencarian dalam periode waktu tertentu.

Menurut Lean Yu [11], data *Google Trends* dapat digunakan sebagai *input* tambahan untuk meningkatkan prediksi konsumsi minyak dengan MAPE sebesar 1.45% dan RMSE 1.7963 menggunakan metode SVR. Lean Yu menyebutkan bahwa terdapat hubungan kointegrasi dan *Granger Causality* antara data yang ingin diprediksi dengan data *Google Trends* menggunakan kata kunci yang berhubungan.

Sedangkan menurut Hongping Hu [2], data *Google Trends* dapat meningkatkan prediksi indeks harga saham. Akurasi *hit ratio* yang didapat menggunakan metode ISCA-BPNN dengan *Google Trends* sebesar 86.81% untuk indeks S&P 500 dan 88.98% untuk indeks DJIA. Hongping Hu mengatakan bahwa data pencarian indeks harga saham mencerminkan perhatian orang pada perubahan pergerakan indeks harga saham itu sendiri.

Dalam tugas akhir ini dilakukan prediksi pergerakan indeks harga saham IHSG dan LQ45 dengan melibatkan data *Google Trends* menggunakan metode *Support Vector Regression* (SVR). Serta melakukan uji Kointegrasi dan uji *Granger Causality* dari data *Google Trends* dan data indeks harga saham. Untuk melihat tingkat akurasi yang didapat dari prediksi indeks harga saham di Indonesia melibatkan data *Google Trends*.

Topik dan Batasannya

Topik masalah pada penelitian tugas akhir ini adalah memprediksi pergerakan indeks Harga Saham di Indonesia dengan melibatkan data *Google Trends*. Dengan mengimplementasikan Uji Kointegrasi dan analisis *Granger Causality Test* untuk mengetahui asosiasi (keterhubungan) antara data indeks harga saham (IHSG dan LQ45) dengan data *Google Trends*. Serta memprediksi pergerakan indeks harga saham (IHSG dan LQ45) dengan melibatkan data *Google Trends* menggunakan metode *Support Vector Regression*.

Batasan Masalah dalam tugas akhir ini adalah sebagai berikut:

- Data yang digunakan untuk diprediksi adalah data penutupan (*closing*) mingguan dari indeks harga saham IHSG dan LQ45 selama 10 tahun terhitung dari tanggal 1 Januari 2009 sampai 1 Oktober 2019 sebanyak 562 data. Sumber data IHSG diperoleh dari situs www.finance.yahoo.com dan data LQ45 diperoleh dari situs www.investing.com.
- Data *Google Trends* yang digunakan adalah data popularitas pencarian dengan kata kunci untuk masing-masing IHSG dan LQ45 selama 10 tahun terhitung dari tanggal 1 Januari 2009 sampai 1 Oktober 2019 sebanyak 562 data. Sumber data *Google Trends* diperoleh dari situs www.trends.google.com.
- Kata kunci *Google Trends* yang digunakan untuk data IHSG adalah : IHSG, IHSG hari ini, dan IHSG saham.
- Kata kunci *Google Trends* yang digunakan untuk data LQ45 adalah : LQ45, LQ45 hari ini, dan LQ45 saham.
- Metode yang digunakan untuk memprediksi pergerakan indeks harga saham (IHSG dan LQ45) dengan melibatkan data *Google Trends* adalah *Support Vector Regression*.

Tujuan

Tujuan dari penelitian Tugas Akhir ini adalah mengetahui cara memprediksi indeks IHSG dan LQ45 dengan melibatkan data *Google Trends* dengan metode *Support Vector Regression*. Serta Memperoleh hasil prediksi model indeks IHSG dan LQ45 dengan melibatkan data *Google Trends* dan tanpa melibatkan data *Google Trends* menggunakan metode *Support Vector Regression*.

Organisasi Tulisan

Selanjutnya pada Bab 2 membahas mengenai studi terkait, meliputi Uji Kointegrasi, *Granger Causality Test*, *Support Vector Regression*, dan Pengukuran Hasil Prediksi. Pada Bab 3 membahas mengenai sistem yang dibangun. Pada Bab 4 membahas mengenai hasil dan analisis pengujian. Pada Bab 5 membahas mengenai kesimpulan dan saran.

2. Studi Terkait

Studi Terkait berdasarkan penelitian yang dilakukan oleh Lean Yu et al, 2019, berjudul “*Online big data-driven oil consumption forecasting with Google trends*” [11]. Pada penelitian ini, peneliti mengatakan bahwa data *Google Trends* dapat digunakan sebagai *input* tambahan untuk meningkatkan prediksi konsumsi minyak dengan melakukan Uji Kointegrasi dan analisis *Granger Causality* pada data konsumsi minyak dengan data *Google Trends* untuk melihat *predictive power* dari *Google Trends*. Serta melakukan uji prediksi dengan

beberapa metode Algoritma. Metode yang memiliki hasil prediksi terbaik adalah metode *Support Vector Regression* (SVR) dengan MAPE 1.45% dan RMSE 1.7963.

Selanjutnya berdasarkan penelitian yang dilakukan oleh Hongping Hu et al, 2018, berjudul “*Predicting the direction of stock markets using optimized neural networks with Google Trend*” [2]. Pada penelitian ini, peneliti mengatakan bahwa data *Google Trends* dapat membantu meningkatkan akurasi prediksi pergerakan indeks harga saham S&P 500 dan DJIA dengan kata kunci *Google Trends* “S&P 500” dan “DJIA”. Metode prediksi yang digunakan adalah *Improved Sine Cosine Back Propagation Neural Networks* (ISCA-BPNN) dengan *hit ratio* sebesar 86.81% untuk prediksi indeks S&P 500 dan 88.98% untuk indeks DJIA.

2.1 Uji Kointegrasi

Uji Kointegrasi merupakan pengujian analisis stasioneritas data untuk melihat keterhubungan jangka panjang antara dua variabel atau lebih. Salah satu metode untuk melakukan uji kointegrasi adalah *Engle-Granger* [12]. Tahapan pada metode *Engle-Granger* adalah membangun residual (error) berdasarkan regresi statis dan residual diuji untuk setiap *unit root* yang ada menggunakan *Augmented Dickey-Fuller* (ADF) atau pengujian serupa [13]. Regresi linear pada stasioneritas ϕ_t dan ψ_t dapat dirumuskan seperti berikut [11]:

$$\phi_t = \phi_0 + \phi_1 \phi_t + \epsilon_t \tag{1}$$

Dimana ϕ_0 merupakan konstanta dan ϕ_1 parameter model linier *time series*. Dilakukan uji stasionaritas pada residual series ϵ_t dengan uji ADF. Apabila ϕ_t terbukti stasioner, maka terdapat hubungan kointegrasi antara ϕ_t dan ψ_t . Pada uji ADF, diberikan data *time series* $\phi_1, \phi_2, \dots, \phi_n$. Uji ADF mempertimbangkan tiga persamaan *autoregresif* bentuk diferensial untuk mendeteksi *unit root*. Yaitu Persamaan tanpa *constant* dan *trend* (2), persamaan dengan *constant* dan tanpa *trend* (3), serta persamaan dengan *constant* dan *trend* (4):

$$\Delta \phi_t = \alpha_1 \phi_{t-1} + \sum_{i=1}^p (\beta_i \Delta \phi_{t-i}) + \epsilon_t \tag{2}$$

$$\Delta \phi_t = \alpha_0 + \alpha_1 \phi_{t-1} + \sum_{i=1}^p (\beta_i \Delta \phi_{t-i}) + \epsilon_t \tag{3}$$

$$\Delta \phi_t = \alpha_0 + \alpha_1 \phi_{t-1} + \sum_{i=1}^p (\beta_i \Delta \phi_{t-i}) + \gamma t + \epsilon_t \tag{4}$$

Dimana t adalah waktu, $\Delta \phi_t$ merupakan selisih dari $\phi_t - \phi_{t-1}$, α_0 konstanta yang disebut *drift*,

α_1 merupakan koefisien *time trend*, β_i adalah koefisien yang menampilkan *unit root*, p adalah *lag order*, $\Delta \phi_{t-i}$ merupakan selisih dari $\phi_{t-i} - \phi_{t-i-1}$, γ merupakan koefisien dari $\Delta \phi_{t-i}$, dan ϵ_t adalah *residual*

term. Persamaan yang digunakan dalam uji ADF adalah persamaan dengan *constant* tanpa *trend* (3), dimana pada perumusan tersebut jika $\alpha_0 = 0$, $\alpha_1 \neq 0$ maka ϕ_t merupakan *random walk* pada *trend*, data tidak stasioner dan terdapat *unit root* pada data. Jika data stasioner, maka $\alpha_0 < 0$. α_0 merupakan *null* hipotesis.

Hubungan yang dilihat adalah hasil *stasionarity test* dari *t-statistics* (*p-value*), *t-statistics* merupakan hasil uji statistik dari ADF, dan *p-value* merupakan probabilitas dari hasil uji statistik. Setelah itu dapat dilihat hasil *stasionarity test* dari data *Google Trends* dengan data indeks harga saham untuk kemudian dapat dilakukan pemilihan data *Google Trends* yang digunakan untuk dilakukan uji *Granger Causality*. Perumusan *t-statistics* dari uji ADF adalah seperti berikut:

$$t_{\alpha_1} = \frac{\hat{\alpha}_1}{SE(\hat{\alpha}_1)} \tag{5}$$

$$t_{\alpha_1}(\phi) = \frac{\hat{\alpha}_1}{\sqrt{SE(\hat{\alpha}_1)}} \tag{6}$$

Dimana SE merupakan *standart error* dari prediksi koefisien α_1 yang menampilkan *unit root* dan n merupakan banyaknya sampel. Apabila data *time series* tidak stasioner pada tingkat data *original* (data sebenarnya), maka dilakukan *differencing* untuk mendapatkan data yang stasioner. Perumusan *differencing* dari data *time series* ϕ_t adalah sebagai berikut:

$$\phi_t = \phi_t - \phi_{t-1} \tag{7}$$

Differencing merupakan selisih data antara data ke t dan $t-1$. Dua data *time series* ϕ_t dan ψ_t terkointegrasi apabila keduanya stasioner pada tingkat stasioneritas data yang sama.

2.2 Granger Causality

Granger Causality adalah pengujian hipotesis statistik untuk menentukan apakah suatu data deret waktu (*time series*) dapat berguna dalam sebuah peramalan (*forecasting*) [14]. Uji *Granger Causality* dapat digunakan dalam melihat efek dari *Google Trends* pada indeks harga saham di Indonesia.

Granger Causality pada stasioner X_t ke Y_t dalam persamaan (1) dapat didefinisikan dalam perumusan berikut [11]:

$$(8) \text{ Untuk } (i = 1, 2, \dots, p) \text{ dan } (j = 1, 2, \dots, q) \text{ maka } \phi_{ij}(h) = \phi_{ij}(h | \phi_{ij}(h-1) - \dots - \phi_{ij}(h-p))$$

Dimana $\phi_{ij}(h | \phi_{ij}(h-1) - \dots - \phi_{ij}(h-p))$ adalah kondisi probabilitas distribusi dari X_t berdasarkan data informasi bivariat

$\{X_{t-1}, Y_{t-1}\}$, dimana $\phi_{ij}(h) = \phi_{ij}(h-1) - \dots - \phi_{ij}(h-p)$ dan $\phi_{ij}(h) = \phi_{ij}(h-1) - \dots - \phi_{ij}(h-p)$. Jika pada $\{X_{t-1}, Y_{t-1}\}$

persamaan (8) termasuk dalam *statistical rejected*, maka dapat dibuktikan bahwa X_t dapat membantu untuk memprediksi Y_t . Lalu *Vector Autoregression* (VAR) dapat digunakan untuk memodelkan hubungan kausalitas:

$$\begin{aligned} X_t &= \alpha_0 + \alpha_1 X_{t-1} + \dots + \alpha_p X_{t-p} + \epsilon_t, \\ Y_t &= \beta_0 + \beta_1 X_{t-1} + \dots + \beta_p X_{t-p} + \gamma_0 + \gamma_1 Y_{t-1} + \dots + \gamma_q Y_{t-q} + \eta_t, \\ X_t &= \alpha_0 + \alpha_1 X_{t-1} + \dots + \alpha_p X_{t-p} + \epsilon_t, \\ Y_t &= \beta_0 + \beta_1 X_{t-1} + \dots + \beta_p X_{t-p} + \gamma_0 + \gamma_1 Y_{t-1} + \dots + \gamma_q Y_{t-q} + \eta_t, \end{aligned}$$

(10)

Pada perumusan (9) dan (10), terdapat *restricted* model (atas) dan *unrestricted* model (bawah). Dimana ϵ_t dan η_t merupakan error dengan *zero means* dan *constant variances*. Dilakukan *joint test* (uji χ^2 atau F^2) untuk menguji signifikansi dari koefisien $\alpha_i (i = 1, \dots, p)$ dan $\beta_j (j = 1, \dots, q)$. Perumusan *F*-

statistic pada uji F adalah sebagai berikut:

$$F\text{-stat} = \frac{ESS_{restricted} - ESS_{unrestricted}}{ESS_{unrestricted} / (n - k)}$$

Dimana ESS merupakan *error* dari *restricted* dan *unrestricted* model X_t mempengaruhi Y_t (9) atau Y_t mempengaruhi X_t (10). Sedangkan q merupakan banyaknya data pada *restricted* model, n merupakan total data, dan k merupakan banyaknya data pada *unrestricted* model. Jika koefisien terbukti pada deviasi *joint* dari nol, maka *Granger Causality* dapat dilakukan dari X_t ke Y_t atau sebaliknya dengan persamaan (9) dan (10). Uji *Granger Causality* dilakukan dengan uji *null* hipotesis H_0 dengan perumusan:

$$H_0: \phi_{12} = 0 \text{ (} X_t \text{ tidak granger terhadap } Y_t \text{)} \tag{12}$$

$$H_0: \phi_{21} = 0 \text{ (} Y_t \text{ tidak granger terhadap } X_t \text{)} \tag{13}$$

Uji *null* hipotesis dapat dilakukan dengan uji *F-test*. *F-test* adalah uji statistik untuk melihat pengaruh variabel bebas terhadap variabel terikat. *Null* hipotesis ditolak jika nilai *p-value* kurang dari nilai signifikan yang ditentukan.

2.3 Support Vector Regression (SVR)

Support Vector Regression (SVR) merupakan metode yang pertama kali diperkenalkan oleh Harris Drucker pada tahun 1996 [15]. SVR merupakan kasus regresi dalam penerapan *Support Vector Machine* (SVM). Dalam kasus regresi SVR dapat mengatasi *overfitting*, *output* yang dikeluarkan adalah bilangan riil ataupun kontinu. Metode SVR dapat menemukan fungsi regresi $f(x) = w \cdot x + b$ yang memiliki nilai deviasi yang besar dari data *training* yang diberikan sehingga dapat memprediksi nilai aktualnya. Bentuk persamaan fungsi regresi dalam SVR [16] adalah sebagai berikut:

Dimana (\cdot) merupakan *dot product*, $w \in X, b \in R$. Regresi pada SVR yaitu mencari persamaan (14) ke data regresi dengan mencari nilai w yang minimal dengan meminimalkan bentuk Euclidean dalam $\|w\|^2$. Dapat dirumuskan sebagai *convex optimization* seperti persamaan berikut:

$$\begin{aligned} & \text{Minimize } \frac{1}{2} \|w\|^2 \\ & \text{Subject to } \{ \langle w, x_i \rangle - b - \xi_i \leq \epsilon \\ & \langle w, x_i \rangle + b + \xi_i \leq \epsilon \end{aligned} \tag{15}$$

Pada persamaan diatas, fungsi regresi $f(x)$ dapat mengaproksimasi titik (x_i, y_i) dalam zona ϵ -*insensitive*. Untuk mengatasi permasalahan *convex optimization* yang *feasible* yaitu semua titik berada

dalam rentang nilai $(\hat{y}_i) \pm \epsilon$ dapat ditambahkan variabel *slack* yaitu ξ_i dan ξ_i^* sebagai batasan *infeasible constrain* dalam optimasi. Dapat dirumuskan sebagai berikut :

$$\begin{aligned}
 & \text{Minimize } \frac{1}{2} \|\hat{y}\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\
 & \text{Subject to } \left\{ \begin{aligned} & \hat{y}_i - \langle \hat{y}_i \cdot \hat{y}_i \rangle - \epsilon \leq \xi_i \\ & \langle \hat{y}_i \cdot \hat{y}_i \rangle + \epsilon - \hat{y}_i \leq \xi_i^* \end{aligned} \right. \quad (16)
 \end{aligned}$$

Konstanta $C > 0$ merupakan *trade off* antara kedataran fungsi dan batasan toleransi kesalahan (*error*) pada model. Besar kecilnya nilai C dapat mempengaruhi pada hasil prediksi model. Jika C terlalu besar maka toleransi error akan semakin kecil. Dengan kata lain jika kesalahan yang memiliki nilai melebihi ϵ akan dikenakan *punishment* yang besar. Jika C terlalu kecil maka toleransi error akan bernilai besar. Pemilihan nilai C harus optimal sehingga hasil prediksi model SVR dapat menjadi lebih akurat. Berikut adalah fungsi pada kerugian *ϵ-insensitive*:

$$|\xi|_i := \begin{cases} 0 & , \text{ jika } |\xi| \leq \epsilon \\ |\xi| - \epsilon & , \text{ jika } |\xi| > \epsilon \end{cases} \quad (17)$$

Masalah pada optimasi dapat diselesaikan pada teori *Lagrangian* sebagai berikut

$$\begin{aligned}
 L &= \frac{1}{2} \|\hat{y}\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\
 & - \sum_{i=1}^n \alpha_i (\epsilon + \xi_i - y_i + \langle \hat{y}_i \cdot \hat{y}_i \rangle + \epsilon) \\
 & - \sum_{i=1}^n \alpha_i^* (\epsilon + \xi_i^* + y_i - \langle \hat{y}_i \cdot \hat{y}_i \rangle - \epsilon) \\
 & - \sum_{i=1}^n (\eta_i \xi_i + \eta_i^* \xi_i^*) \quad (18)
 \end{aligned}$$

$$\text{Untuk } : \eta_i, \eta_i^*, \alpha_i, \alpha_i^* \geq 0$$

Pada persamaan (18), L adalah *Lagrangian*, sedangkan $\eta_i, \eta_i^*, \alpha_i, \alpha_i^*$ merupakan *Lagrangian multiplier*.

Dibutuhkan turunan parsial *Lagrangian* terhadap variabel *primal* w, b, ξ_i , dan ξ_i^* yang dapat dijabarkan sebagai berikut:

$$\frac{\partial L}{\partial w} = \sum_{i=1}^n (\alpha_i^* - \alpha_i) \hat{y}_i = 0 \quad (19)$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \hat{y}_i = 0 \quad (20)$$

$$\frac{\partial L}{\partial \xi_i} = C - \alpha_i - \eta_i = 0 \quad (21)$$

Hasil turunan parsial *Lagrangian* (19), (20), (21) dapat disubstitusi ke dalam persamaan (18) sehingga didapatkan perumusan pada optimasi *dual* dengan perumusan sebagai berikut:

$$\begin{aligned}
 & \text{Maximize } \left\{ \frac{1}{2} \sum_{i=1}^n (\alpha_i - \alpha_i^*) (\alpha_i - \alpha_i^*) \langle \hat{y}_i \cdot \hat{y}_i \rangle \right. \\
 & \left. - \sum_{i=1}^n \alpha_i y_i + \sum_{i=1}^n \alpha_i^* y_i \right\} \quad (22) \\
 & \text{Subject to } \sum_{i=1}^n (\alpha_i - \alpha_i^*) \langle \hat{y}_i \cdot \hat{y}_i \rangle + \sum_{i=1}^n \alpha_i y_i - \sum_{i=1}^n \alpha_i^* y_i = 0
 \end{aligned}$$

Perumusan subjek optimasi dual dapat diturunkan menjadi:

Sehingga didapatkan :

$$\phi = \sum_{\phi=1}^n (\alpha_{\phi} - \alpha_{\phi}^*) \phi_{\phi} \quad (23)$$

$$f(x) = \langle \phi \cdot \phi \rangle + b$$

$$f(x) = \sum_{\phi=1}^n (\alpha_{\phi} - \alpha_{\phi}^*) \langle \phi_{\phi} \cdot \phi \rangle + b \quad (24)$$

Untuk kasus dimensi data yang besar dapat digunakan fungsi *kernel* sehingga didapatkan perumusan:

$$f(\phi, \alpha_{\phi}, \alpha_{\phi}^*) = \sum_{\phi=1}^n (\alpha_{\phi} - \alpha_{\phi}^*) K(\phi_{\phi} \cdot \phi) + b \quad (25)$$

Dalam perumusan $f(x)$ diatas, ϕ_{ϕ} merupakan *Support Vector* (SV) yang menyimpan informasi dari data training dan berada pada $\phi(\phi) \pm \epsilon$. Dimana α_{ϕ} dan α_{ϕ}^* merupakan variabel *Lagrangian multiplier*.

Sedangkan $K(\phi_{\phi} \cdot \phi) = (f(\phi_{\phi})f(\phi))$ merupakan fungsi *kernel* dimana merupakan *linear mapping dot product non linear mapping*. Banyak fungsi *kernel* yang dapat digunakan dalam SVR, contohnya adalah fungsi *kernel* linier, polinomial, *Gaussian radial basis*, dan *sigmoid*. Nilai b dapat dirumuskan dengan prinsip Karush Khun Tucker (KKT) sebagai berikut:

$$\begin{aligned} \alpha_{\phi} &= \alpha_{\phi} - \langle \phi \cdot \phi_{\phi} \rangle - \epsilon \\ \alpha_{\phi}^* &= \alpha_{\phi} - \langle \phi \cdot \phi_{\phi} \rangle + \epsilon \end{aligned}$$

$$\text{dimana } 0 \leq \alpha_{\phi}, \alpha_0^* \leq C \tag{26}$$

Model SVR yang digunakan dalam tugas akhir ini adalah model SVR dengan fungsi *kernel* yaitu terdapat pada perumusan (25). Fungsi *kernel* yang digunakan adalah *Gaussian radial basis* [17].

(27) Dengan ϕ merupakan parameter dan fungsi *kernel*. *Gaussian radial basis* merupakan fungsi *kernel* yang dapat mengatasi dimensi data yang besar.

2.4 Partial Auto-Correlation Function

Partial Auto Correlation Function (PACF) merupakan perhitungan hubungan korelasi secara parsial antara data ϕ_t dan data ϕ_{t-k} sampai *lag* k [18]. Sebagai contoh diberikan data ϕ_t dan ϕ_{t-1} yang bergantung pada ϕ_{t-2} . Adapun perumusan PACF adalah korelasi antara ϕ_t dan ϕ_{t-1} setelah menghilangkan pengaruh ϕ_{t-2} . Adapun perumusan PACF adalah sebagai berikut.

$$\phi_{\phi\phi\phi\phi}(\phi) = \frac{(\phi_{\phi\phi\phi\phi}(\phi) - \phi_{\phi\phi\phi\phi}(\phi-1)\phi_{\phi\phi\phi\phi}(\phi-2) - \dots - \phi_{\phi\phi\phi\phi}(\phi-k+1)\phi_{\phi\phi\phi\phi}(\phi-k))}{\phi_{\phi\phi\phi\phi}(\phi-k+1)} \times \phi_{\phi\phi\phi\phi}(\phi-k+1) \tag{28}$$

Dimana ϕ_t merupakan data time series pada waktu ke i, ϕ_{t-1} adalah data time series pada waktu ke i-1, dan k adalah jumlah *lag*.

2.5 Pengukuran Hasil Prediksi

Pengukuran hasil akurasi prediksi dalam kasus regresi bisa dilakukan dengan menggunakan beberapa ukuran error. Seperti *Root Mean Squared Error* (RMSE), *Mean Absolute Percentage Error* (MAPE), dan *Hit Ratio*. RMSE dapat menentukan seberapa besaran nilai rata-rata dari jumlah kuadrat kesalahan prediksi. RMSE berfungsi untuk melihat besar kesalahan dari model prediksi. Perumusan RMSE dapat dirumuskan sebagai berikut:

$$\phi_{\phi\phi\phi\phi} = \sqrt{\frac{\sum_{t=1}^N (\phi_t - \phi_t^*)^2}{N}} \tag{29}$$

Dimana ϕ_t adalah hasil prediksi, ϕ_t^* adalah data aktual, dan n adalah total data. Sedangkan MAPE merupakan total dari persentase rata-rata nilai *error* dengan data aktual. MAPE dapat dirumuskan sebagai berikut:

$$\phi_{\phi\phi\phi\phi} = \frac{\sum_{t=1}^n \frac{|\phi_t - \phi_t^*|}{\phi_t^*}}{n} \times 100 \tag{30}$$

Dimana ϕ_t adalah hasil prediksi dan $\phi_t \neq 0$, ϕ_t^* adalah data aktual, dan n adalah jumlah data.

Sedangkan untuk menentukan ukuran besar pergerakan prediksi indeks harga saham dapat menggunakan rumus *Hit Ratio* [2] dengan rumus sebagai berikut:

$$\phi_{\phi\phi\phi\phi} = \frac{\sum_{t=1}^n \phi_t}{n} \times 100 \tag{31}$$

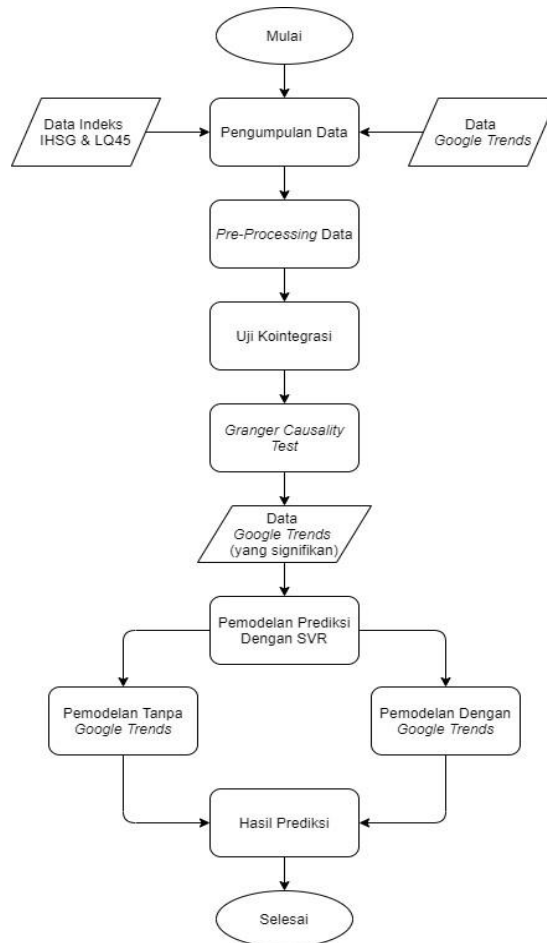
$$\phi_t := \begin{cases} 1, & (\phi_{t+1} - \phi_t)(\phi_{t+1}^* - \phi_t^*) > 0 \\ 0, & \text{lainnya} \end{cases} \tag{32}$$

Dimana ϕ_t dapat dijabarkan pada perumusan (32). Sedangkan ϕ_t adalah data aktual pada waktu ke i, ϕ_t adalah data prediksi pada waktu ke i, dan n adalah banyaknya *testing* sampel.

3. Sistem Yang Dibangun

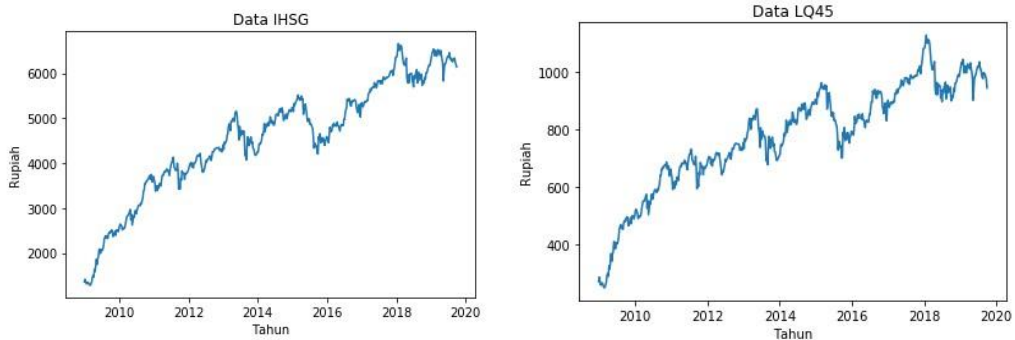
Sistem yang dibangun dalam Tugas Akhir ini adalah pemodelan untuk memprediksi pergerakan indeks harga saham dengan melibatkan data *Google Trends* menggunakan metode *Support Vector Regression* (SVR). *Output* dari sistem yang dibuat adalah nilai akurasi prediksi pergerakan indeks harga saham di masa mendatang. Untuk dapat melibatkan data *Google Trends* sebagai data *input* prediksi indeks harga saham, dibutuhkan uji keterhubungan antara data *Google Trends* dengan data indeks harga saham. Uji keterhubungan

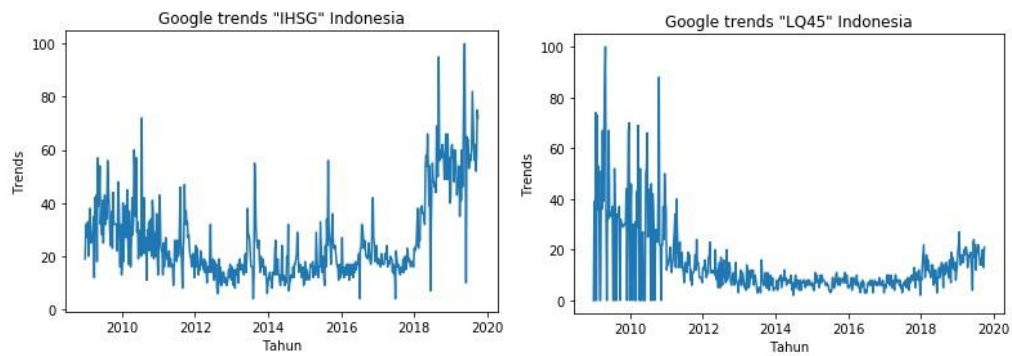
data dilakukan dengan cara uji Kointegrasi dan uji *Granger Causality*. Setelah melakukan uji keterhubungan data, dilakukan pemilihan data *Google Trends* yang memiliki keterhubungan data dengan data indeks harga saham untuk digunakan sebagai *input* prediksi.



Gambar 3.1 Flowchart Sistem Prediksi Pergerakan Indeks Harga Saham dengan *Google Trends*
 Keterangan:

- Pengumpulan Data
 Data IHSG, LQ45, dan *Google Trends* yang digunakan adalah data mingguan sebanyak 562 data terhitung dari 1 Januari 2009 sampai 1 Oktober 2019. Untuk data indeks harga saham digunakan data penutupan harga mingguan. Data *Google Trends* yang digunakan adalah data dengan *keyword* IHSG, IHSG hari ini, IHSG saham, LQ45, LQ45 hari ini, dan LQ45 saham. Data *Google Trends* yang digunakan adalah sebanyak 12 data dengan filter *Worldwide* dan filter *Indonesia*. Berikut merupakan gambaran grafik data yang digunakan :





Gambar 3.2 Grafik Data

Gambar grafik data lebih lengkap dapat dilihat pada lampiran.

- *Pre-Processing Data*
Pre-Processing data yang dilakukan adalah menormalkan data *Google Trends* mingguan yang didapat per 5 tahun terhitung dari 1 Januari 2009 sampai 1 Oktober 2019. Serta melakukan *cleansing* data indeks harga saham IHSG dan LQ45.
- Uji Kointegrasi
 Dilakukan uji Kointegrasi antara data index harga saham dengan *Google Trends* untuk melihat keterhubungan data.
- *Granger Causality Test*
 Dilakukan uji *Granger Causality* untuk melihat seberapa besar pengaruh data *Google Trends* sebagai prediktor dari prediksi indeks harga saham.
- Mendapatkan Data *Google Trends* yang Signifikan
 Setelah melakukan uji *Granger Causality*, didapatkan data *Google Trends* yang memiliki keterhubungan signifikan dengan data indeks harga saham untuk pemodelan prediksi.
- Pemodelan SVR
 Dilakukan uji PACF untuk menentukan *lag* prediktor indeks harga saham yang digunakan pada pemodelan. Digunakan dua metode untuk memprediksi indeks harga saham IHSG dan LQ45.
 1. Metode pertama adalah pemodelan tanpa *Google Trends*, dengan menggunakan harga penutupan indeks harga saham minggu ini tanpa melibatkan data *Google Trends* sebagai prediktor untuk memprediksi harga penutupan indeks harga saham minggu depan. Penggunaan metode ini berdasarkan hasil PACF data indeks IHSG dan LQ45.
 2. Metode kedua adalah pemodelan dengan *Google Trends*, menggunakan harga penutupan indeks harga saham minggu ini dan data *Google Trends* sebagai prediktor untuk memprediksi harga penutupan indeks harga saham minggu depan. Penggunaan metode ini berdasarkan hasil PACF data indeks IHSG dan LQ45 dan hasil *Granger Causality Test* dari data *Google Trends*.

Parameter SVR C, epsilon, dan gamma yang digunakan adalah parameter dengan nilai 0.1, 1, dan 10. Dilakukan 3 skenario pemodelan SVR untuk pengujian:

- Skenario 1, terdapat 562 data, kemudian data yang kita gunakan disusun sehingga terdapat 500 data prediktor dan 4 data untuk diprediksi. Sebagai contoh kita memiliki data *time series* $\diamond_{t+1}, \diamond_{t+2}, \dots, \diamond_{t+500}$ sebagai prediktor untuk memprediksi data $\diamond_{t+501}, \diamond_{t+502}, \dots, \diamond_{t+504}$ waktu t berjalan dari 0 sampai 58. Hasil akurasi dari 58 kali pemodelan dirata-ratakan sehingga didapatkan hasil rata-rata akurasi prediksi untuk masing-masing parameter.
- Skenario 2, mengambil 523 data dari 1 Januari 2009 sampai 31 Desember 2018 sebagai data untuk *training model* dan 39 data dari 1 Januari 2019 sampai 1 Oktober 2019 sebagai data uji. Prediksi yang dilakukan adalah prediksi 1 waktu kedepan.
- Skenario 3, mengambil 523 data dari 1 Januari 2009 sampai 31 Desember 2018 sebagai data untuk *training model* dan 39 data dari 1 Januari 2019 sampai 1 Oktober 2019 sebagai data testing. Prediksi yang dilakukan adalah prediksi 3 waktu kedepan.
- Hasil Prediksi
 Hasil uji akurasi prediksi yang digunakan adalah hasil uji MAPE, RMSE, dan *hit ratio*. Lalu hasil uji akurasi prediksi dengan melibatkan data *Google Trends* dan tanpa melibatkan data *Google Trends* dibandingkan untuk mendapatkan kesimpulan hasil.

4. Evaluasi

4.1 Hasil Pengujian

4.1.1 Uji Kointegrasi

Data *Google Trends* yang dapat dilakukan Uji Kointegrasi dengan data indeks harga saham adalah data yang memiliki kemiripan stasioneritas. Apabila data indeks harga saham stasioner pada Tingkat Data Original, maka data *Google Trends* yang dapat dilakukan Uji Kointegrasi adalah data yang stasioner pada Tingkat Data Original.

Pertama, dilakukan Uji Stasioneritas menggunakan ADF test. Dapat dilihat pada Tabel 1, bahwa Indeks IHSG, *Google trends* 'IHSG' (Indonesia), *Google trends* 'ihsg saham' (Indonesia), *Google trends* 'ihsg hari ini' (Indonesia), *Google trends* 'IHSG' (Worldwide), *Google trends* 'ihsg saham' (Worldwide), dan *Google trends* 'ihsg hari ini' (Worldwide) stasioner pada Tingkat Data *Differencing* dengan signifikan 10% level. Hasil uji stasioneritas menyatakan bahwa data *Google Trends* IHSG yang stasioner pada Tingkat Data *Differencing* dapat dilakukan Uji Kointegrasi dan *Granger Causality Test* dengan data Indeks IHSG.

Tabel 1. Uji Stasioneritas Data IHSG

Hasil tes stasioneritas dalam <i>t-statistics</i> (<i>p-values</i>). Tulisan dengan tebal dan garis bawah merupakan hasil dengan signifikan 10% level.		
<i>Time series</i>	Tingkat Data Original	Tingkat Data <i>Differencing</i>
Indeks IHSG	-2.263042 (0.184165)	-26.821884 (<u>0.000000</u>)
<i>Google trends</i> 'IHSG' (Indonesia)	0.022511 (0.960355)	-11.576299 (<u>0.000000</u>)
<i>Google trends</i> 'ihsg saham' (Indonesia)	-1.394090 (0.585106)	-9.294942 (<u>0.000000</u>)
<i>Google trends</i> 'ihsg hari ini' (Indonesia)	-0.114382 (0.947987)	-17.006956 (<u>0.000000</u>)
<i>Google trends</i> 'IHSG' (Worldwide)	0.479454 (0.984235)	-5.420665 (<u>0.000003</u>)
<i>Google trends</i> 'ihsg saham' (Worldwide)	2.989787 (1.000000)	-7.942395 (<u>0.000000</u>)
<i>Google trends</i> 'ihsg hari ini' (Worldwide)	0.420531 (0.982247)	-14.799918 (<u>0.000000</u>)

Pada Tabel 2, Indeks LQ45, *Google trends* 'lq45 saham' (Indonesia), *Google trends* 'LQ45' (Worldwide), *Google trends* 'lq45 saham' (Worldwide), dan *Google trends* 'lq45 hari ini' (Worldwide) stasioner pada Tingkat Data *Differencing* dengan signifikan 10% level. Hasil uji stasioneritas menyatakan bahwa data *Google Trends* LQ45 yang stasioner pada Tingkat Data *Differencing* dapat dilakukan Uji Kointegrasi dan *Granger Causality Test* dengan data Indeks LQ45.

Tabel 2. Uji Stasioneritas Data LQ45

Hasil tes stasioneritas dalam <i>t-statistics</i> (<i>p-values</i>). Tulisan dengan tebal dan garis bawah merupakan hasil dengan signifikan 10% level.		
<i>Time series</i>	Tingkat Data Original	Tingkat Data <i>Differencing</i>
Indeks LQ45	-2.562917 (0.100906)	-28.220458 (<u>0.000000</u>)
<i>Google trends</i> 'LQ45' (Indonesia)	-4.318908 (<u>0.000412</u>)	-
<i>Google trends</i> 'lq45 saham' (Indonesia)	-1.952615 (0.307728)	-9.018127 (<u>0.000000</u>)
<i>Google trends</i> 'lq45 hari ini' (Indonesia)	-5.905306 (<u>0.000000</u>)	-
<i>Google trends</i> 'LQ45' (Worldwide)	0.829093 (0.992092)	-9.567178 (<u>0.000000</u>)
<i>Google trends</i> 'lq45 saham' (Worldwide)	-0.043337 (0.954798)	-9.109495 (<u>0.000000</u>)
<i>Google trends</i> 'lq45 hari ini' (Worldwide)	-2.390461 (0.144447)	-13.159197 (<u>0.000000</u>)

Kedua, dilakukan Uji Kointegrasi antara data indeks IHSG dan data *Google Trends*. Dapat dilihat pada Tabel 3, bahwa Uji Kointegrasi antara data indeks IHSG dengan data *Google trends* 'IHSG' (Indonesia), *Google trends* 'ihsg saham' (Indonesia), *Google trends* 'ihsg hari ini' (Indonesia), *Google trends* 'IHSG' (Worldwide), *Google trends* 'ihsg saham' (Worldwide), dan *Google trends* 'ihsg hari ini' (Worldwide) memiliki hasil signifikan pada 10% level dan dapat dilakukan uji *Granger Causality*.

Tabel 3. Uji Kointegrasi Data IHSG

Hasil uji kointegrasi dalam hal <i>t</i> -statistics (<i>p</i> -values)	
<i>Google Trend</i>	Tingkat Data <i>Differencing</i>
<i>Google trends</i> 'IHSG' (Indonesia)	-27.489219 (0.000000)
<i>Google trends</i> 'ihsg saham' (Indonesia)	-26.944009 (0.000000)
<i>Google trends</i> 'ihsg hari ini' (Indonesia)	-27.140635 (0.000000)
<i>Google trends</i> 'IHSG' (<i>Worldwide</i>)	-27.609419 (0.000000)
<i>Google trends</i> 'ihsg saham' (<i>Worldwide</i>)	-26.915670 (0.000000)
<i>Google trends</i> 'ihsg hari ini' (<i>Worldwide</i>)	-27.145910 (0.000000)

Dapat dilihat pada Tabel 4, bahwa Uji Kointegrasi antara data indeks LQ45 dengan data *Google trends* 'lq45 saham' (Indonesia), *Google trends* 'LQ45' (*Worldwide*), *Google trends* 'lq45 saham' (*Worldwide*), dan *Google trends* 'lq45 hari ini' (*Worldwide*) memiliki hasil signifikan pada 10% level dan dapat dilakukan uji *Granger Causality*.

Tabel 4. Uji Kointegrasi Data LQ45

Hasil uji kointegrasi dalam hal <i>t</i> -statistics (<i>p</i> -values)	
<i>Google Trend</i>	Tingkat Data <i>Differencing</i>
<i>Google trends</i> 'lq45 saham' (Indonesia)	-28.232189 (0.000000)
<i>Google trends</i> 'LQ45' (<i>Worldwide</i>)	-28.290389 (0.000000)
<i>Google trends</i> 'lq45 saham' (<i>Worldwide</i>)	-28.286466 (0.000000)
<i>Google trends</i> 'lq45 hari ini' (<i>Worldwide</i>)	-28.251515 (0.000000)

4.1.2 *Granger Causality Test*

Hasil uji *Granger Causality* data *Google Trends* sebagai prediktor untuk data indeks IHSG pada Tabel 5 menunjukkan bahwa *Google Trends* dengan keyword 'ihsg hari ini' filter *Worldwide* memiliki hasil signifikan 10% level pada lag 3. *Google Trends* 'ihsg hari ini' (*Worldwide*) lag 3 dapat digunakan pada pengujian sebagai prediktor tambahan untuk pemodelan prediksi indeks IHSG dengan melibatkan data *Google Trends*.

Tabel 5. Uji Granger Causality Data IHSG

Hasil <i>Granger Causality Test</i> data <i>Google Trends</i> dengan indeks IHSG						
	<i>Lags</i>					
	1	2	3	4	5	6
Panel A	H0 : <i>Google trend</i> of 'IHSG' (Indonesia) does not <i>Granger cause</i> index ihsg					
<i>F-Stat</i>	0.064	0.0256	0.3315	0.6422	0.592	0.452
<i>p-value</i>	0.8004	0.9748	0.8026	0.6326	0.7061	0.8437
Panel B	H0 : <i>Google trend</i> of 'ihsg hari ini' (Indonesia) does not <i>Granger cause</i> index ihsg					
<i>F-Stat</i>	0.2775	1.5444	1.6739	1.4512	1.1077	1.2963
<i>p-value</i>	0.5985	0.2144	0.1715	0.2158	0.3551	0.2569
Panel C	H0 : <i>Google trend</i> of 'ihsg saham' (Indonesia) does not <i>Granger cause</i> index ihsg					
<i>F-Stat</i>	0.7537	0.5174	0.3549	0.2851	0.2379	0.2576
<i>p-value</i>	0.3857	0.5963	0.7856	0.8877	0.9457	0.9562
Panel D	H0 : <i>Google trend</i> of 'IHSG' (<i>Worldwide</i>) does not <i>Granger cause</i> index ihsg					
<i>F-Stat</i>	0.9663	0.874	1.757	1.5714	1.5504	1.2214
<i>p-value</i>	0.326	0.4178	0.1543	0.1805	0.1725	0.2935
Panel E	H0 : <i>Google trend</i> of 'ihsg hari ini' (<i>Worldwide</i>) does not <i>Granger cause</i> index ihsg					
<i>F-Stat</i>	0.0187	0.0635	2.3209	1.918	1.5382	1.2766
<i>p-value</i>	0.8913	0.9385	0.0743	0.106	0.1761	0.2661

Panel F	H0 : Google trend of 'ihsg saham' (Worldwide) does not Granger cause index ihsg					
F-Stat	0.4025	0.197	0.6123	0.5311	0.9961	0.8916
p-value	0.5261	0.8212	0.6072	0.7129	0.4194	0.5006

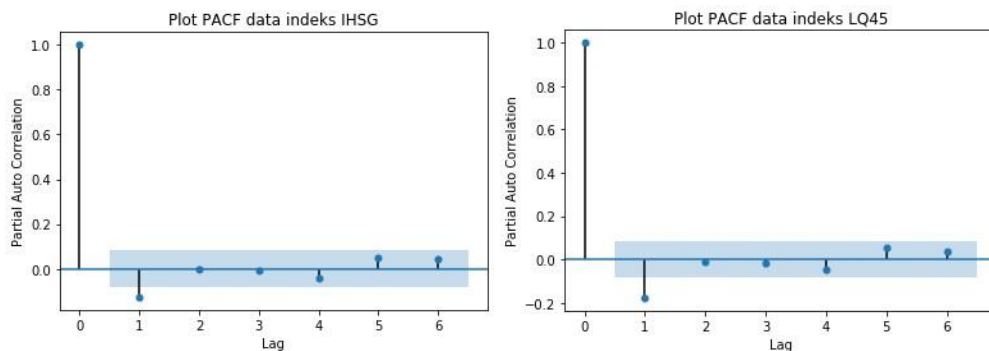
Hasil uji *Granger Causality* data *Google Trends* sebagai prediktor untuk data indeks LQ45 pada Tabel 6 menunjukkan bahwa *Google Trends* dengan keyword 'LQ45' filter *Worldwide* memiliki hasil signifikan 10% level pada lag 3, 4, 5, dan 6. *Google Trends* 'LQ45' (*Worldwide*) lag 3, 4, 5, dapat digunakan pada pengujian sebagai prediktor tambahan untuk pemodelan prediksi indeks LQ45 dengan melibatkan data *Google Trends*.

Tabel 6. Uji Granger Causality Data LQ45

Hasil <i>Granger Causality Test</i> data <i>Google Trends</i> dengan indeks LQ45						
	Lags					
	1	2	3	4	5	6
Panel A	H0 : Google trend of 'lq45 saham' (Indonesia) does not Granger cause index lq45					
F-Stat	0.1253	0.5966	0.69	0.5149	0.437	0.6345
p-value	0.7235	0.551	0.5584	0.7248	0.8228	0.7027
Panel B	H0 : Google trend of 'LQ45' (Worldwide) does not Granger cause index lq45					
F-Stat	0.0853	0.4199	3.2824	2.6301	2.4166	2.4404
p-value	0.7704	0.6573	0.0206	0.0336	0.035	0.0245
Panel C	H0 : Google trend of 'lq45 hari ini' (Worldwide) does not Granger cause index lq45					
F-Stat	0.1462	0.3835	0.2626	0.4912	1.0839	0.8945
p-value	0.7024	0.6817	0.8523	0.7422	0.3682	0.4985
Panel D	H0 : Google trend of 'lq45 saham' (Worldwide) does not Granger cause index lq45					
F-Stat	0.0788	0.2372	0.3145	0.3147	0.1936	0.4536
p-value	0.7791	0.7889	0.8149	0.8682	0.965	0.8425

4.1.3 Pemodelan SVR

Sesuai dengan hasil *Granger Causality Test*, pemodelan prediksi melibatkan *Google Trends* untuk IHSG menggunakan keyword 'ihsg hari ini' (*Worldwide*) dengan lag 3 dan untuk LQ45 menggunakan keyword 'LQ45' (*Worldwide*) dengan lag 3, 4, 5, dan 6. Dilakukan uji skenario 1, 2, dan 3 dengan pemodelan melibatkan data *Google Trends* dan tanpa melibatkan data *Google Trends*. Berikut merupakan hasil dari pemodelan prediksi IHSG skenario 1. Dilakukan uji PACF untuk menentukan lag data indeks harga saham pada pemodelan prediksi yang dibuat.



Gambar 4.1 Plot PACF Data IHSG dan LQ45

Pada gambar 4.1, sumbu x merupakan lag order, dan sumbu y merupakan nilai *Partial Autocorrelation*. Garis biru merupakan *significant threshold* atau ambang batas hasil yang signifikan. Lag dikatakan signifikan jika hasilnya berada diluar *significant threshold*. Lag yang termasuk signifikan untuk indeks IHSG dan LQ45 berdasarkan plot PACF adalah data indeks harga sama pada lag 1. Dengan nilai pacf sebesar -0.1266 untuk IHSG, dan -0.1773 untuk LQ45. Oleh karena itu, pada pemodelan prediksi tanpa *Google Trends* untuk IHSG dan LQ45 menggunakan harga penutupan indeks harga saham pada lag 1 sebagai prediktor. Untuk pemodelan prediksi IHSG dengan *Google Trends*

digunakan indeks IHSG lag 1 dan *Google Trends* 'ihsg hari ini' (*Worldwide*) lag 3 sebagai prediktor. Untuk pemodelan prediksi LQ45 dengan *Google Trends* digunakan indeks LQ45 lag 1 dan *Google Trends* 'LQ45' (*Worldwide*) lag 3, 4, 5, dan 6 sebagai prediktor. Berikut merupakan hasil pemodelan prediksi:

Tabel 7. Hasil Pemodelan Prediksi IHSG Skenario 1

No	IHSG Skenario 1 Parameter	Pemodelan Tanpa <i>Google Trends</i>			Pemodelan Dengan <i>Google Trends</i>		
		MAPE	RMSE	Hit Ratio	MAPE	RMSE	Hit Ratio
1	C = 0.1 , epsilon = 0.1, gamma = 0.1	1.7852	124.3620	58.0460	1.7855	124.3962	58.0460
2	C = 0.1 , epsilon = 0.1, gamma = 1	1.7855	124.3878	58.0460	1.7862	124.4419	58.0460
3	C = 0.1 , epsilon = 0.1, gamma = 10	1.7851	124.3554	58.0460	1.7860	124.4245	58.0460
4	C= 1 , epsilon = 0.1, gamma = 0.1	1.7884	124.5681	58.0460	1.7904	124.7534	58.0460
5	C= 1 , epsilon = 0.1, gamma = 1	1.7898	124.6894	58.0460	1.7894	124.6719	58.0460
6	C= 1 , epsilon = 0.1, gamma = 10	1.7896	124.7104	58.0460	1.7878	124.5508	58.0460
7	C= 10 , epsilon = 0.1, gamma = 0.1	1.8410	128.2923	58.0460	1.8149	126.4666	59.7701
8	C= 10 , epsilon = 0.1, gamma = 1	1.8407	128.1786	51.1494	1.8013	125.5552	58.0460
9	C= 10 , epsilon = 0.1, gamma = 10	1.8108	126.5125	58.0460	1.7886	124.6311	58.0460
10	C = 0.1 , epsilon = 1, gamma = 0.1	1.7871	124.4855	58.0460	1.7874	124.5224	58.0460
11	C = 0.1 , epsilon = 1, gamma = 1	1.7875	124.5245	58.0460	1.7878	124.5512	58.0460
12	C = 0.1 , epsilon = 1, gamma = 10	1.7870	124.4894	58.0460	1.7877	124.5437	58.0460
13	C= 1 , epsilon = 1, gamma = 0.1	1.7871	124.4710	58.0460	1.7899	124.7057	58.0460
14	C= 1 , epsilon = 1, gamma = 1	1.7871	124.4741	58.0460	1.7891	124.6498	58.0460
15	C= 1 , epsilon = 1, gamma = 10	1.7880	124.5814	58.0460	1.7870	124.4959	58.0460
16	C= 10 , epsilon = 1, gamma = 0.1	1.8390	128.1785	58.0460	1.8174	126.6294	59.7701
17	C= 10 , epsilon = 1, gamma = 1	1.8390	128.0537	51.1494	1.8017	125.5806	58.0460
18	C= 10 , epsilon = 1, gamma = 10	1.8125	126.6236	58.0460	1.7883	124.6037	58.0460
19	C = 0.1 , epsilon = 10, gamma = 0.1	1.7748	123.6269	58.0460	1.7784	123.8782	58.0460
20	C = 0.1 , epsilon = 10, gamma = 1	1.7756	123.6757	58.0460	1.7784	123.8778	58.0460
21	C = 0.1 , epsilon = 10, gamma = 10	1.7750	123.6359	58.0460	1.7782	123.8630	58.0460
22	C= 1 , epsilon = 10, gamma = 0.1	1.7809	123.9944	58.0460	1.7826	124.1626	58.0460
23	C= 1 , epsilon = 10, gamma = 1	1.7830	124.1714	58.0460	1.7787	123.9057	58.0460
24	C= 1 , epsilon = 10, gamma = 10	1.7807	124.0344	58.0460	1.7771	123.7871	58.0460
25	C= 10 , epsilon = 10, gamma = 0.1	1.8430	128.4785	54.5977	1.8330	127.6406	59.7701
26	C= 10 , epsilon = 10, gamma = 1	1.8390	128.1230	52.8736	1.8039	125.6874	58.0460
27	C= 10 , epsilon = 10, gamma = 10	1.8296	127.5683	58.0460	1.7890	124.6434	58.0460
Rata-Rata		1.8004	125.4536	57.2158	1.7913	124.8007	58.2375

Berikut merupakan hasil dari pemodelan prediksi LQ45 skenario 1.

Tabel 8. Hasil Pemodelan Prediksi LQ45 Skenario 1

No	LQ45 Skenario 1 Parameter	Pemodelan Tanpa <i>Google Trends</i>			Pemodelan Dengan <i>Google Trends</i>		
		MAPE	RMSE	Hit Ratio	MAPE	RMSE	Hit Ratio
1	C = 0.1 , epsilon = 0.1, gamma = 0.1	2.2753	25.0785	56.3218	2.2789	25.0894	56.3218
2	C = 0.1 , epsilon = 0.1, gamma = 1	2.2784	25.1004	56.3218	2.2788	25.0891	56.3218
3	C = 0.1 , epsilon = 0.1, gamma = 10	2.2781	25.0873	56.3218	2.2788	25.0891	56.3218
4	C= 1 , epsilon = 0.1, gamma = 0.1	2.2934	25.3902	52.2989	2.2787	25.0852	56.3218
5	C= 1 , epsilon = 0.1, gamma = 1	2.3501	25.9518	51.7241	2.2780	25.0776	56.3218
6	C= 1 , epsilon = 0.1, gamma = 10	2.3057	25.4895	52.8736	2.2780	25.0776	56.3218

7	C= 10 , epsilon = 0.1, gamma = 0.1	2.4671	27.2308	56.3218	<u>2.2815</u>	<u>25.1322</u>	56.3218
8	C= 10 , epsilon = 0.1, gamma = 1	2.6402	29.0854	49.4253	<u>2.2805</u>	<u>25.1191</u>	<u>56.3218</u>
9	C= 10 , epsilon = 0.1, gamma = 10	2.6641	29.2240	43.6782	<u>2.2805</u>	<u>25.1191</u>	<u>56.3218</u>
10	C = 0.1 , epsilon = 1, gamma = 0.1	2.2794	25.1163	56.3218	<u>2.2755</u>	<u>25.0534</u>	56.3218
11	C = 0.1 , epsilon = 1, gamma = 1	2.2840	25.1583	56.3218	<u>2.2755</u>	<u>25.0530</u>	56.3218
12	C = 0.1 , epsilon = 1, gamma = 10	2.2791	25.1043	56.3218	<u>2.2755</u>	<u>25.0530</u>	56.3218
13	C= 1 , epsilon = 1, gamma = 0.1	2.3112	25.5340	51.7241	<u>2.2776</u>	<u>25.0787</u>	<u>56.3218</u>
14	C= 1 , epsilon = 1, gamma = 1	2.3437	25.8870	54.5977	<u>2.2774</u>	<u>25.0754</u>	<u>56.3218</u>
15	C= 1 , epsilon = 1, gamma = 10	2.3170	25.6141	52.8736	<u>2.2774</u>	<u>25.0754</u>	<u>56.3218</u>
16	C= 10 , epsilon = 1, gamma = 0.1	2.4740	27.2961	57.4713	<u>2.2816</u>	<u>25.1312</u>	56.3218
17	C= 10 , epsilon = 1, gamma = 1	2.6063	28.7408	51.1494	<u>2.2808</u>	<u>25.1216</u>	<u>56.3218</u>
18	C= 10 , epsilon = 1, gamma = 10	2.6497	29.1087	47.1264	<u>2.2808</u>	<u>25.1216</u>	<u>56.3218</u>
19	C = 0.1 , epsilon = 10, gamma = 0.1	2.2922	25.2506	56.3218	<u>2.2826</u>	<u>25.1235</u>	56.3218
20	C = 0.1 , epsilon = 10, gamma = 1	2.2880	25.1992	56.3218	<u>2.2827</u>	<u>25.1239</u>	56.3218
21	C = 0.1 , epsilon = 10, gamma = 10	2.2851	25.1636	56.3218	<u>2.2827</u>	<u>25.1239</u>	56.3218
22	C= 1 , epsilon = 10, gamma = 0.1	2.3505	25.9623	51.1494	<u>2.2879</u>	<u>25.1606</u>	<u>56.3218</u>
23	C= 1 , epsilon = 10, gamma = 1	2.3296	25.7430	54.0230	<u>2.2878</u>	<u>25.1602</u>	<u>56.3218</u>
24	C= 1 , epsilon = 10, gamma = 10	2.3135	25.5481	54.5977	<u>2.2878</u>	<u>25.1602</u>	<u>56.3218</u>
25	C= 10 , epsilon = 10, gamma = 0.1	2.5685	28.3261	53.4483	<u>2.2918</u>	<u>25.1869</u>	<u>56.3218</u>
26	C= 10 , epsilon = 10, gamma = 1	2.4871	27.5777	50.0000	<u>2.2916</u>	<u>25.1839</u>	<u>56.3218</u>
27	C= 10 , epsilon = 10, gamma = 10	2.4195	26.8477	56.3218	<u>2.2915</u>	<u>25.1830</u>	56.3218
Rata-Rata		2.3863	26.3265	53.6186	<u>2.2816</u>	<u>25.1129</u>	<u>56.3218</u>

Tabel hasil pemodelan prediksi skenario 2 dan 3 terlampir pada lampiran.

4.2 Analisis Hasil Pengujian

4.2.1 Analisis Hasil Pengujian Skenario 1

Tabel 9. Hasil Prediksi Terbaik IHSG Skenario 1

IHSG Skenario 1	Parameter	MAPE	RMSE	Hit Ratio
Pemodelan dengan <i>Google Trends</i>	C = 0.1 , epsilon = 10, gamma = 0.1	1.7771	123.7871	58.0460
Pemodelan dengan <i>Google Trends</i>	C= 10 , epsilon = 0.1, gamma = 0.1	1.8149	126.4666	<u>59.7701</u>
Pemodelan tanpa <i>Google Trends</i>	C= 1 , epsilon = 10, gamma = 10	<u>1.7748</u>	<u>123.6269</u>	58.0460

Tabel 10. Rata-Rata Hasil Prediksi IHSG Skenario 1

Rata-Rata Pengujian IHSG Skenario 1	MAPE	RMSE	Hit Ratio
Pemodelan dengan <i>Google Trends</i>	<u>1.7913</u>	<u>124.8007</u>	<u>58.2375</u>
Pemodelan tanpa <i>Google Trends</i>	1.8004	125.4536	57.2158

Pada Tabel 9 pengujian pemodelan prediksi IHSG Skenario 1, hasil MAPE dan RMSE terbaik adalah dengan pemodelan tanpa *Google Trends* menggunakan parameter C = 1, Epsilon = 10, Gamma = 10 dengan hasil MAPE 1.7748 dan RMSE 123.6269. Hasil uji *Hit Ratio* terbaik adalah 59.7701 menggunakan pemodelan dengan *Google Trends* dan parameter C= 10 , epsilon = 0.1, gamma = 0.1.

Pada table 10 hasil rata-rata pengujian IHSG Skenario 1 dengan *Google Trends* memiliki hasil MAPE, RMSE, dan *Hit Ratio* yang lebih baik. Namun, hasil prediksi terbaik adalah dengan pemodelan tanpa *Google Trends*.

Tabel 11. Hasil Prediksi Terbaik LQ45 Skenario 1

LQ45 Skenario 1	Parameter	MAPE	RMSE	Hit Ratio
Pemodelan dengan <i>Google Trends</i>	C = 0.1 , epsilon = 1, gamma = 1	2.2755	<u>25.0530</u>	56.3218

Pemodelan tanpa <i>Google Trends</i>	C = 0.1 , epsilon = 0.1, gamma = 0.1	<u>2.2753</u>	25.0785	56.3218
Pemodelan tanpa <i>Google Trends</i>	C= 10 , epsilon = 1, gamma = 0.1	2.4740	27.2961	<u>57.4173</u>

Tabel 12. Rata-Rata Hasil Prediksi LQ45 Skenario 1

Rata-Rata Pengujian LQ45 Skenario 1	MAPE	RMSE	Hit Ratio
Pemodelan dengan <i>Google Trends</i>	<u>2.2816</u>	<u>25.1129</u>	<u>56.3218</u>
Pemodelan tanpa <i>Google Trends</i>	2.3863	26.3265	53.6186

Pada Tabel 11 pengujian pemodelan prediksi LQ45 Skenario 1, hasil MAPE terbaik adalah dengan pemodelan tanpa *Google Trends* menggunakan parameter C = 1, Epsilon = 10, Gamma = 10 dengan hasil MAPE 2.2753, dan hasil RMSE terbaik adalah dengan pemodelan menggunakan *Google Trends* dengan parameter C = 0.1, Epsilon = 1, Gamma = 1 dengan hasil RMSE 123.6269. Hasil uji *Hit Ratio* terbaik adalah 57.4173 menggunakan pemodelan tanpa *Google Trends* dan parameter C= 10 , epsilon = 1, gamma = 0.1.

Pada Tabel 12 hasil rata-rata pengujian LQ45 Skenario 1 dengan *Google Trends* memiliki hasil MAPE, RMSE, dan *Hit Ratio* yang lebih baik. Namun, hasil prediksi terbaik untuk MAPE terkecil adalah dengan pemodelan tanpa *Google Trends*. Sedangkan untuk RMSE terkecil adalah dengan pemodelan dengan *Google Trends*.

4.2.2 Analisis Hasil Pengujian Skenario 2

Tabel 13. Hasil Prediksi Terbaik IHSG Skenario 2

IHSG Skenario 2	Parameter	MAPE	RMSE	Hit Ratio
Pemodelan dengan <i>Google Trends</i>	C= 1 , epsilon = 10, gamma = 0.1	4.4889	342.9022	<u>50</u>
Pemodelan tanpa <i>Google Trends</i>	C = 1 , epsilon = 10, gamma = 0.1	<u>4.4341</u>	<u>338.2187</u>	<u>50</u>

Tabel 14. Rata-Rata Hasil Prediksi IHSG Skenario 2

Rata-Rata Pengujian IHSG Skenario 2	MAPE	RMSE	Hit Ratio
Pemodelan dengan <i>Google Trends</i>	<u>4.9169</u>	<u>377.8594</u>	<u>50.0000</u>
Pemodelan tanpa <i>Google Trends</i>	4.9305	379.5716	49.6101

Pada Tabel 13 pengujian pemodelan prediksi IHSG Skenario 2, hasil MAPE dan RMSE terbaik adalah dengan pemodelan tanpa *Google Trends* menggunakan parameter C = 1, Epsilon = 10, Gamma = 0.1 dengan hasil MAPE 4.4341 dan RMSE 338.2187. Hasil uji *Hit Ratio* pada pengujian pemodelan prediksi IHSG skenario 2 memiliki hasil yang tidak signifikan berbeda.

Pada Tabel 14 hasil rata-rata pengujian IHSG Skenario 2 dengan *Google Trends* memiliki hasil MAPE, RMSE, dan *Hit Ratio* yang lebih baik. Namun, hasil prediksi terbaik adalah dengan pemodelan tanpa *Google Trends*.

Tabel 15. Hasil Prediksi Terbaik LQ45 Skenario 2

LQ45 Skenario 2	Parameter	MAPE	RMSE	Hit Ratio
Pemodelan dengan <i>Google Trends</i>	C= 10 , epsilon = 10, gamma = 1	3.7163	47.3271	44.7368
Pemodelan tanpa <i>Google Trends</i>	C = 10 , epsilon = 10, gamma = 10	<u>3.6976</u>	<u>46.0152</u>	42.1053
Pemodelan tanpa <i>Google Trends</i>	C= 1 , epsilon = 0.1, gamma = 0.1	6.3892	77.9048	<u>50.0000</u>

Tabel 16. Rata-Rata Hasil Prediksi LQ45 Skenario 2

Rata-Rata Pengujian LQ45 Skenario 2	MAPE	RMSE	Hit Ratio
Pemodelan dengan <i>Google Trends</i>	<u>4.7803</u>	<u>59.4512</u>	<u>44.7368</u>
Pemodelan tanpa <i>Google Trends</i>	5.7850	70.6542	42.2027

Pada Tabel 15 pengujian pemodelan prediksi LQ45 Skenario 2, hasil MAPE dan RMSE terbaik adalah dengan pemodelan tanpa *Google Trends* menggunakan parameter C = 10, Epsilon = 10, Gamma

= 10 dengan hasil MAPE 3.6976 dan RMSE 46.0152. Hasil uji *Hit Ratio* terbaik adalah 50.0000 menggunakan pemodelan tanpa *Google Trends* dan parameter $C=1$, $\epsilon=0.1$, $\gamma=0.1$.

Pada Tabel 16 hasil rata-rata pengujian LQ45 Skenario 2 dengan *Google Trends* memiliki hasil MAPE, RMSE, dan *Hit Ratio* yang lebih baik. Namun, hasil prediksi terbaik adalah dengan pemodelan tanpa *Google Trends*.

4.2.3 Analisis Hasil Pengujian Skenario 3

Tabel 17. Hasil Prediksi Terbaik IHSG Skenario 3

IHSG Skenario 3	Parameter	MAPE	RMSE	<i>Hit Ratio</i>
Pemodelan dengan <i>Google Trends</i>	$C=1$, $\epsilon=10$, $\gamma=0.1$	1.5763	110.4858	51.3889
Pemodelan dengan <i>Google Trends</i>	$C=10$, $\epsilon=10$, $\gamma=0.1$	1.6002	111.9229	52.7778
Pemodelan tanpa <i>Google Trends</i>	$C=1$, $\epsilon=10$, $\gamma=0.1$	1.5765	110.5510	51.3889
Pemodelan tanpa <i>Google Trends</i>	$C=10$, $\epsilon=10$, $\gamma=10$	1.5772	110.5195	52.7778

Tabel 18. Rata-Rata Hasil Prediksi IHSG Skenario 3

Rata-Rata Pengujian IHSG Skenario 3	MAPE	RMSE	<i>Hit Ratio</i>
Pemodelan dengan <i>Google Trends</i>	1.5893	111.3682	51.4403
Pemodelan tanpa <i>Google Trends</i>	1.5954	111.7721	50.6687

Pada Tabel 17 pengujian pemodelan prediksi IHSG Skenario 3, hasil MAPE dan RMSE terbaik adalah dengan pemodelan dengan melibatkan data *Google Trends* menggunakan parameter $C=1$, $\epsilon=10$, $\gamma=0.1$ dengan hasil MAPE 1.5763 dan RMSE 110.4858. Hasil uji *Hit Ratio* pada pengujian pemodelan prediksi IHSG skenario 3 memiliki hasil yang tidak signifikan berbeda.

Pada Tabel 18 hasil rata-rata pengujian IHSG Skenario 3 dengan *Google Trends* memiliki hasil MAPE, RMSE, dan *Hit Ratio* yang lebih baik.

Tabel 19. Hasil Prediksi Terbaik LQ45 Skenario 3

LQ45 Skenario3	Parameter	MAPE	RMSE	<i>Hit Ratio</i>
Pemodelan dengan <i>Google Trends</i>	$C=10$, $\epsilon=10$, $\gamma=0.1$	2.0605	22.7098	45.8333
Pemodelan tanpa <i>Google Trends</i>	$C=0.1$, $\epsilon=10$, $\gamma=10$	2.0668	22.8020	45.8333
Pemodelan tanpa <i>Google Trends</i>	$C=10$, $\epsilon=1$, $\gamma=0.1$	2.1214	23.3251	56.9444

Tabel 20. Rata-Rata Hasil Prediksi LQ45 Skenario 3

Rata-Rata Pengujian LQ45 Skenario 3	MAPE	RMSE	<i>Hit Ratio</i>
Pemodelan dengan <i>Google Trends</i>	2.0715	22.8559	45.8333
Pemodelan tanpa <i>Google Trends</i>	2.1218	23.4062	47.4280

Pada Tabel 19 pengujian pemodelan prediksi LQ45 Skenario 3, hasil MAPE dan RMSE terbaik adalah dengan pemodelan dengan melibatkan data *Google Trends* menggunakan parameter $C=10$, $\epsilon=10$, $\gamma=0.1$ dengan hasil MAPE 2.0605 dan RMSE 22.7098. Hasil uji *Hit Ratio* terbaik adalah 56.9444 menggunakan pemodelan tanpa *Google Trends* dan parameter $C=10$, $\epsilon=1$, $\gamma=0.1$.

Pada Tabel 20 hasil rata-rata pengujian IHSG Skenario 3 dengan *Google Trends* memiliki hasil MAPE dan RMSE yang lebih baik. Serta untuk rata-rata *Hit Ratio*, pemodelan tanpa *Google Trends* memiliki hasil lebih baik.

5. Kesimpulan

Kesimpulan yang bisa diambil dalam tugas akhir ini adalah:

1. Pemodelan prediksi indeks IHSG terbaik dengan melibatkan data *Google Trends* menggunakan keyword 'ihsg hari ini' (*Worldwide*) terdapat pada pengujian skenario 3 dengan hasil MAPE 1.5763% dan RMSE 110.4858. Serta pengujian skenario 1 dengan *Hit Ratio* 59.7701%. Pemodelan prediksi IHSG terbaik tanpa melibatkan *Google Trends* terdapat pada pengujian skenario 3 dengan hasil MAPE 1.5765% dan RMSE 110.5195. Serta pengujian skenario 1 dengan *Hit Ratio* 58.0460%.

2. Pemodelan prediksi indeks LQ45 terbaik dengan melibatkan data *Google Trends* menggunakan keyword 'LQ45' (*Worldwide*) terdapat pada pengujian skenario 3 dengan hasil MAPE 2.0605% dan RMSE 22.7098. Serta pengujian skenario 1 dengan Hit Ratio 56.3218%. Pemodelan prediksi LQ45 terbaik tanpa melibatkan *Google Trends* terdapat pada pengujian skenario 3 dengan hasil MAPE 2.0668% dan RMSE 22.8020. Serta pengujian skenario 1 dengan Hit Ratio 57.4173%.
3. Pemodelan prediksi indeks IHSG dan LQ45 dengan melibatkan data *Google Trends* memiliki akurasi lebih baik dibandingkan pemodelan tanpa melibatkan data *Google Trends*. Namun nilai akurasi prediksi kedua model tidak signifikan berbeda.

Saran yang ingin disampaikan untuk pengembangan lebih lanjut adalah:

1. Model prediksi dengan melibatkan data *Google Trends* masih dapat diteliti lebih lanjut dengan mencari parameter yang optimal agar mendapatkan hasil akurasi lebih baik.
2. Menggunakan keyword *Google Trends* lain yang sekiranya berhubungan untuk meningkatkan akurasi prediksi model.

Daftar Pustaka

- [1] Bursa Efek Indonesia. 2018. Informasi Umum Saham. [Online] Available at: <https://www.idx.co.id/produk/saham/> [Accessed 28 November 2019].
- [2] Hu H, Tang L, Zhang S, dan Wang H. 2018. *Predicting the direction of stock markets using optimized neural networks with Google Trends*. ScienceDirect. *Neurocomputing*. vol. 285 188-195
- [3] Ariyo A A, Adewumi A O, dan Ayo C K. 2014. *Stock Price Prediction Using the ARIMA Model*. UKSim-AMSS 16th International Conference on Computer Modelling and Simulation. IEEE. Durban.
- [4] W. Huang, Y. Nakamori, dan S. -Y. Wang. 2004. *Forecasting stock market movement direction with support vector machine*. ScienceDirect. *Computers & Operations Research*. vol. 32(10) 2513-2522
- [5] Y. Kara, M. A. Boyacioglu, dan . Ö. K. Baykan. 2011. *Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange*. ScienceDirect. *Expert Systems with Applications*. vol. 38(5) 5311-5319
- [6] L. Khaidem, S. Saha, dan S. R. Dey. 2016. *Predicting the direction of stock market prices using random forest*. arXiv. *Mathematical Finance*. 1-20
- [7] M. Qiu dan Y. Song. 2016. *Predicting the Direction of Stock Market Index Movement Using an Optimized Artificial Neural Network Model*. PLOS ONE. vol. 11(5) e0155133
- [8] T. B. Trafalis dan H. Ince. 2000. *Support Vector Machine for Regression and Applications to Financial Forecasting*. *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks*. IJCNN 2000. *Neural Computing: New Challenges and Perspectives for the New Millennium*.
- [9] C.-Y. Yeh, C.-W. Huang, dan S.-J. Lee. 2011. *A Multiple-Kernel Support Vector Regression Approach for Stock Market Price Forecasting*. ScienceDirect. *Expert Systems with Applications*. vol. 38(3) 2177-2186
- [10] X.-d. Zhang, A. Li, dan R. Pan. 2016. *Stock Trend prediction Based on a New Status Box Method and AdaBoost Probabilistic Support Vector Machine*. ScienceDirect. *Applied Soft Computing*. vol. 49 385-398
- [11] L. Yu, Y. Zhao, L. Tang, dan Z. Yang. 2019. *Online big data-driven oil consumption forecasting with Google trends*. ScienceDirect. *International Journal Forecasting*. vol. 35 213-223
- [12] Engle, R. F., & Granger, C. W. J. 1987. *Co-Integration and Error Correction: Representation, Estimation, and Testing*. *Econometrica*, vol. 55(2)
- [13] K. S. Lai dan M. Lai. 1991. *A Cointegration Test for Market Efficiency*. *Journal of Futures Markets*, vol. 11(5) 567-575
- [14] C. Hiemstra dan J. D. Jones. 1994. *Testing for Linear and Nonlinear Granger Causality in the Stock Price-*

Volume Relation. The Journal of Finance. vol. 49(5) 1639-1664

- [15] H. Drucker, C. J. Burges, L. Kaufman, A. Smola, dan V. 1997. *Support Vector Regression Machines. Advances in Neural Information Processing Systems.* 155-161
- [16] D. Basak, S. Pal, D. C. Patranabis. 2007. *Support Vector Regression. Neural Information Processing.* vol. 11(10)
- [17] J. P. Vert, K. Tsuda, dan B. Schölkopf. 2004. *A Primer on Kernel Methods. Computational Biology.*
- [18] *Towards Data Science.* 2019. *Understanding Partial Auto-Correlation.* [Online] Available at: <https://towardsdatascience.com/understanding-partial-auto-correlation-fa39271146ac> [Accessed 04 April 2020].