

**PEMODELAN REGRESI SEMIPARAMETRIK DENGAN PENDEKATAN DERET  
FOURIER**  
(Studi Kasus: Pengaruh Indeks *Dow Jones* dan *BI Rate* Terhadap Indeks Harga Saham Gabungan)

Laili Rahma Khairunnisa<sup>1</sup>, Alan Prahutama, S.Si., M.Si.<sup>2</sup>, Dr. Rukun Santoso, M.Si.<sup>3</sup>  
<sup>1,2,3</sup>Departemen Statistika FSM Universitas Diponegoro  
Email : [alan.prahutama@gmail.com](mailto:alan.prahutama@gmail.com)

**ABSTRACT**

The Composite Stock Price Index (CSPI) is a composite index all of types of shares listed on the stock exchange and their movements indicate conditions that occur in the capital market. CSPI is influenced by macroeconomic factors and foreign exchange index. *Dow Jones Industrial Average* has a linear relationship with CSPI and *BI Rate* has a repeated relationship with CSPI, so the method is used semiparametric regression with the Fourier series approach. Estimators in semiparametric regression with Fourier series approach were obtained by the *Ordinary Least Square* (OLS) method. This study uses monthly data which is divided into *in sample* data and *out sample* data. Semiparametric regression modelling with Fourier series approach is done by determining the optimal K value which results in a minimum *General Cross Validation* (GCV) value. In this study, semiparametric regression model with Fourier series approach formed by the optimal K value is 13 and GCV is 2826122. The results of the evaluation of the accuracy of the model performance and forecasting obtained the coefficient of determination is 0,9226, *Mean Absolute Percentage Error* (MAPE) data *in sample* 3,8154% and data *out sample* is 8,4782% which shows that the model obtained has a very accurate performance.

**Keywords:** Composite Stock Price Index (CSPI), Semiparametric Regression, Fourier Series, OLS, GCV

**1. PENDAHULUAN**

Analisis regresi merupakan salah satu metode yang banyak dikaji oleh peneliti statistika. Secara umum analisis regresi adalah suatu metode statistika yang digunakan untuk mengetahui pola hubungan antara variabel respon dengan variabel prediktor. Tujuan utama dalam analisis regresi adalah untuk mengestimasi kurva regresi. Terdapat tiga bentuk pendekatan untuk mengestimasi kurva regresi yaitu pendekatan regresi parametrik, pendekatan regresi nonparametrik dan pendekatan regresi semiparametrik.

Pendekatan regresi parametrik digunakan untuk pemodelan hubungan antar variabel jika bentuk hubungan fungsionalnya diketahui atau diasumsikan mengikuti fungsi tertentu yaitu linier, kuadratik, kubik, eksponensial dan lain sebagainya (Gujarati, 2004) Pendekatan nonparametrik digunakan untuk menjelaskan bentuk hubungan fungsional yang tidak diketahui antara variabel respon dan prediktor. Dalam pendekatan regresi nonparametrik tidak diperlukan asumsi seperti pada regresi parametrik. Regresi nonparametrik memiliki kemampuan dalam mencari bentuk pola kurva regresi yang tidak diketahui bentuknya. Pendekatan regresi semiparametrik digunakan apabila dalam kasus tertentu dalam pemodelan terdapat komponen parametrik, yaitu hubungan antara variabel respon dan prediktor mengikuti pola tertentu atau bentuk fungsinya diketahui dan

komponen nonparametrik, yaitu hubungan antara variabel respon dan prediktor tidak diketahui bentuk polanya atau bentuk fungsinya tidak diketahui (Budiantara, 2014).

Salah satu metode untuk mengestimasi model regresi dengan pendekatan nonparametrik maupun semiparametrik yang paling populer adalah deret Fourier. Deret Fourier merupakan polinomial trigonometri yang mempunyai fleksibilitas. Deret Fourier baik digunakan untuk menjelaskan kurva yang menunjukkan gelombang sinus dan cosinus. Data yang sesuai untuk dianalisis dengan menggunakan deret Fourier adalah apabila data yang diselidiki tidak diketahui polanya dan ada kecenderungan pola data yang berulang (Bilodeau, 1992).

Salah satu bentuk dari berbagai macam kegiatan perekonomian adalah pasar modal. Pasar modal (*capital market*) merupakan tempat bagi investor maupun emiten menjual dan membeli instrumen investasi seperti saham, obligasi, reksadana dan lain-lain pada berbagai sekuritas dengan harapan memperoleh imbalan (*return*). Indeks Harga Saham Gabungan (IHSG) merupakan cerminan dari kegiatan pasar modal secara umum. Menurut Wijaya (2015) pergerakan IHSG dipengaruhi oleh berbagai faktor baik internal maupun eksternal. Pengaruh eksternal yang mempengaruhi IHSG seperti pergerakan indeks saham luar negeri dipercaya telah menjadi faktor dominan yang mempengaruhi IHSG seperti Indeks *Dow Jones* (DJIA), Indeks Nikkei 225, Indeks Hang Seng (HSI) dan lain-lain. Sedangkan faktor internal lebih dipengaruhi oleh peristiwa-peristiwa dalam negeri seperti inflasi, tingkat suku bunga dan nilai tukar Rupiah terhadap Dollar Amerika Serikat. Berdasarkan studi awal penelitian, peneliti membuat *scatterplot* beberapa faktor yang mempengaruhi IHSG, yaitu indeks *Dow Jones* dan *BI Rate*. Dilihat dari *scatterplot* terdapat faktor yang memiliki pola tertentu (komponen parametrik) yaitu indeks *Dow Jones* dan terdapat faktor yang tidak memiliki pola tertentu, acak dan cenderung berulang (komponen parametrik) yaitu *BI Rate*. Karena menggunakan dua faktor dengan pola tertentu (linier) dan tak tentu untuk mendapatkan model terbaik digunakan metode regresi semiparametrik. Pada penelitian ini peneliti menerapkan pendekatan deret Fourier. Masalah dalam penelitian ini dibatasi pada jumlah data yang digunakan yaitu sebanyak 108 data dari bulan Juli 2010 sampai dengan Juni 2019. Tujuan dalam penelitian ini adalah mendentukan pemodelan IHSG dengan regresi semiparametrik deret Fourier.

## 2. TINJAUAN PUSTAKA

### 2.1 Regresi Parametrik

Menurut Hardle (1994) analisis regresi parametrik mengasumsikan bahwa bentuk kurva regresinya sudah diketahui.

Menurut Montgomery dan Runger (2003), model regresi linier dapat dirumuskan seperti pada persamaan (1):

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \varepsilon_i, \quad i = 1, 2, \dots, n; j = 1, 2, \dots, k \quad (1)$$

Jika ditulis dalam bentuk matriks, persamaan (1) dapat ditulis sebagai berikut:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

### 2.2 Estimasi Model Regresi Parametrik

Menurut Gujarati (2004) metode yang digunakan dalam mengestimasi parameter dalam model regresi adalah dengan meminimumkan jumlah kuadrat error dengan *Ordinary Least Square* (OLS). Berdasarkan persamaan (2) diperoleh estimasi  $\boldsymbol{\beta}$  sebagai berikut:

$$\mathbf{L} = \sum_{i=1}^n \varepsilon_i^2 = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = \mathbf{y}^T \mathbf{y} - 2\boldsymbol{\beta}^T \mathbf{X}^T \mathbf{y} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X} \boldsymbol{\beta}$$

dengan syarat bahwa:

$$\frac{\partial L}{\partial \beta} \Big|_{\beta=\hat{\beta}} = \mathbf{0}, \text{ maka diperoleh } -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X} \hat{\beta} = \mathbf{0}$$

$$\mathbf{X}^T \mathbf{X} \hat{\beta} = \mathbf{X}^T \mathbf{y}$$

Sehingga penaksir kuadrat terkecil dari  $\beta$  dapat ditulis sebagai persamaan (3):

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \tag{3}$$

### 2.3 Uji Signifikansi Model

Uji ini dilakukan untuk menentukan keberadaan hubungan linier antara variabel respon Y dengan X dan biasanya dilakukan sebagai prasyarat dalam regresi linier.

Hipotesis:

$H_0 : \beta_1 = 0$  (Tidak terdapat hubungan linier antara variabel X dengan variabel Y)

$H_1 : \beta_1 \neq 0$  (Terdapat hubungan linier antara variabel X dengan variabel Y)

Statistik Uji:

$$F_{hitung} = \frac{RKR}{RKS}$$

$H_0$  ditolak jika  $F_{hitung} > F_{\alpha;1;n-2}$ . Penolakan  $H_0$  menunjukkan bahwa model sesuai atau terdapat hubungan linier antara variabel respon Y dengan variabel prediktor X dalam model.

Prosedur pengujian dapat disajikan dalam tabel *Analysis of Variance* (ANOVA) berikut:

**Tabel 1.** *Analysis of Variance* (ANOVA) untuk Regresi Linier

Sumber Variasi	Derajat kebebasan	Jumlah Kuadrat	Rataan Kuadrat	$F_{hitung}$	$F_{tabel}$
Regresi	1	JKR	RKR	$\frac{RKR}{RKS}$	$F_{\alpha;1;n-2}$
Sesatan	n - 2	JKS	RKS		
Total	n - 1	JKT			

dimana:

$$JKR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

$$JKS = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$JKT = \sum_{i=1}^n (y_i - \bar{y})^2$$

$$RKR = \frac{JKR}{1}$$

$$RKS = \frac{JKS}{n-2}$$

## 2.4 Regresi Nonparametrik

Diberikan data  $(y_i, t_i)$ ,  $i = 1, 2, \dots, n$  dan pola hubungan antara  $y_i$  dan  $t_i$  tidak diketahui maka model regresi nonparametrik secara umum dirumuskan seperti pada persamaan (4):

$$y_i = g(t_i) + \varepsilon_i \quad (4)$$

dengan  $y_i$  adalah variabel respon,  $g(t_i)$  merupakan fungsi regresi yang tidak diketahui bentuknya dan  $t_i$  merupakan variabel prediktor serta  $\varepsilon_i$  diasumsikan  $\varepsilon_i \sim N(0, \sigma^2)$ .

## 2.5 Regresi Nonparametrik dengan Pendekatan Deret Fourier

Deret Fourier merupakan polinomial trigonometri yang mempunyai fleksibilitas yang tinggi. Deret Fourier baik digunakan untuk menjelaskan kurva yang menunjukkan gelombang sinus dan cosinus. Deret fourier digunakan apabila pola data yang terbentuk menunjukkan pola data yang berulang. Dalam regresi nonparametrik dengan pendekatan Fourier selain menggunakan kombinasi aditif fungsi sinus dan cosinus, dapat juga digunakan kombinasi aditif fungsi linier dari fungsi sinus atau cosinus (Suparti *et.al*, 2018). Bilodeau (1992) memberikan memberikan fungsi deret fourier dengan pendekatan kombinasi aditif fungsi linier dan cosinus seperti pada persamaan (5):

$$T(t) = \frac{1}{2}\alpha_0 + \gamma t + \sum_{k=1}^K \alpha_k \cos kt \quad (5)$$

Jika diberikan data sebanyak  $n$   $(t_i, y_i)$ ,  $i = 1, 2, \dots, n$  dengan  $g(t_i)$  merupakan kurva yang tidak diketahui bentuknya dan  $g(t_i)$  didekati dengan menggunakan deret Fourier maka terbentuk persamaan (6):

$$g(t_i) = \frac{1}{2}\alpha_0 + \gamma t_i + \sum_{k=1}^K \alpha_k \cos kt_i \quad (6)$$

Secara umum model regresi nonparametrik pendekatan deret Fourier dengan  $q$  variabel prediktor pada pengamatan ke- $i$  dapat ditulis sebagai persamaan (7):

$$y_i = \sum_{l=1}^q g_l(t_{il}) + \varepsilon_i \quad (7)$$

dengan  $i = 1, 2, \dots, n$  dan  $l = 1, 2, \dots, q$

## 2.6 Regresi Semiparametrik

Regresi semiparametrik merupakan gabungan antara regresi parametrik dan regresi nonparametrik. Diberikan bentuk model regresi semiparametrik dengan komponen parametrik yang mempunyai pola linier seperti pada persamaan (8):

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + g(t_i) + \varepsilon_i \quad (8)$$

dengan  $i = 1, 2, \dots, n$ ;  $j = 1, 2, \dots, k$

Jika ditulis dalam bentuk matriks, persamaan (8) dapat ditulis sebagai persamaan (9):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{g}(\mathbf{t}) + \boldsymbol{\varepsilon} \quad (9)$$

## 2.7 Regresi Semiparametrik dengan Pendekatan Deret Fourier

Model (9) terdiri dari dua komponen yang disebut parametrik yaitu  $\mathbf{X}\boldsymbol{\beta}$  dan komponen yang disebut nonparametrik yaitu  $\mathbf{g}(\mathbf{t})$ , dapat dituliskan sebagai persamaan (10):

$$\mathbf{y}^* = \mathbf{g}(\mathbf{t}) + \boldsymbol{\varepsilon} \quad (10)$$

dengan  $\mathbf{y}^* = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$

Model deret Fourier pada persamaan (6) disubstitusikan ke persamaan (10), dengan asumsi bahwa  $\mathbf{g}(t)$  adalah fungsi aditif maka dapat dituliskan dalam bentuk matriks sebagai persamaan (11):

$$\mathbf{y}^* = \mathbf{C}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \quad (11)$$

dimana  $\mathbf{y}^* = [y_1^* \ y_2^* \ \dots \ y_n^*]^T$

$$\mathbf{C} = \begin{bmatrix} 1 & t_{11} & \cos t_{11} & \cos 2t_{11} & \dots & \cos K t_{11} & \dots & t_{1q} & \cos t_{1q} & \cos 2t_{1q} & \dots & \cos K t_{1q} \\ 1 & t_{21} & \cos t_{21} & \cos 2t_{21} & \dots & \cos K t_{21} & \dots & t_{2q} & \cos t_{2q} & \cos 2t_{2q} & \dots & \cos K t_{2q} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_{n1} & \cos t_{n1} & \cos 2t_{n1} & \dots & \cos K t_{n1} & \dots & t_{nq} & \cos t_{nq} & \cos 2t_{nq} & \dots & \cos K t_{nq} \end{bmatrix}$$

$$\boldsymbol{\theta} = [\emptyset \ \gamma_1 \ \alpha_{11} \ \dots \ \alpha_{1K} \ \dots \ \gamma_q \ \alpha_{q1} \ \dots \ \alpha_{qK}]^T$$

$$\emptyset = \frac{1}{2} \sum_{l=1}^q \alpha_{0l}$$

Estimasi untuk parameter  $\boldsymbol{\theta}$  dapat dicari menggunakan metode OLS (*Ordinary Least Square*) sebagai berikut:

$$\psi(\boldsymbol{\theta}) = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = (\mathbf{y}^* - \mathbf{C}\boldsymbol{\theta})^T (\mathbf{y}^* - \mathbf{C}\boldsymbol{\theta}) = \mathbf{y}^{*T} \mathbf{y}^* - 2\boldsymbol{\theta}^T \mathbf{C}^T \mathbf{y}^* + \boldsymbol{\theta}^T \mathbf{C}^T \mathbf{C} \boldsymbol{\theta}$$

Penaksir kuadrat terkecil harus memenuhi:

$$\left. \frac{\partial \psi(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \mathbf{0}$$

$$-2\mathbf{C}^T \mathbf{y}^* + 2\mathbf{C}^T \mathbf{C} \hat{\boldsymbol{\theta}} = 0$$

$$\mathbf{C}^T \mathbf{C} \hat{\boldsymbol{\theta}} = \mathbf{C}^T \mathbf{y}^*$$

$$\hat{\boldsymbol{\theta}} = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{y}^*$$

Berdasarkan persamaan (10) dan (11), dapat dinyatakan ke dalam persamaan (12):

$$\mathbf{g}(t) = \mathbf{C}\boldsymbol{\theta} \quad (12)$$

$$\hat{\mathbf{g}}(t) = \mathbf{C}\hat{\boldsymbol{\theta}}$$

$$\hat{\mathbf{g}}(t) = \mathbf{C}(\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{y}^*$$

$$\hat{\mathbf{g}}(t) = \mathbf{D}\mathbf{y}^* \quad (13)$$

Berdasarkan persamaan (10) dan (13) dapat ditulis sebagai persamaan (14):

$$\hat{\mathbf{g}}(t) = \mathbf{D}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \quad (14)$$

$$\mathbf{y} - \mathbf{X}\boldsymbol{\beta} = \mathbf{D}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + \boldsymbol{\varepsilon}$$

$$\mathbf{y} - \mathbf{D}\mathbf{y} = \mathbf{X}\boldsymbol{\beta} - \mathbf{D}\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$(\mathbf{I} - \mathbf{D})\mathbf{y} = (\mathbf{I} - \mathbf{D})\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (15)$$

dengan  $\mathbf{I}$  adalah matriks identitas.

Estimasi untuk parameter  $\boldsymbol{\beta}$  dapat diperoleh dengan menggunakan metode OLS. sebagai berikut:

$$\xi(\boldsymbol{\beta}) = ((\mathbf{I} - \mathbf{D})\mathbf{y} - (\mathbf{I} - \mathbf{D})\mathbf{X}\boldsymbol{\beta})^T ((\mathbf{I} - \mathbf{D})\mathbf{y} - (\mathbf{I} - \mathbf{D})\mathbf{X}\boldsymbol{\beta})$$

$$\begin{aligned}
&= (\mathbf{y}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{y}) - (\mathbf{y}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{X} \boldsymbol{\beta}) - \\
&\quad (\boldsymbol{\beta}^T \mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{y}) + (\boldsymbol{\beta}^T \mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{X} \boldsymbol{\beta}) \\
\xi(\boldsymbol{\beta}) &= (\mathbf{y}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{y}) - (\mathbf{y}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{X} \boldsymbol{\beta})^T - \\
&\quad (\boldsymbol{\beta}^T \mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{y}) + (\boldsymbol{\beta}^T \mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{X} \boldsymbol{\beta}) \\
&= (\mathbf{y}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{y}) - 2(\boldsymbol{\beta}^T \mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{y}) + \\
&\quad (\boldsymbol{\beta}^T \mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{X} \boldsymbol{\beta})
\end{aligned}$$

$$\left. \frac{\partial \xi(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \right|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} = \mathbf{0}$$

$$-2(\mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{y}) + 2(\mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{X} \hat{\boldsymbol{\beta}}) = \mathbf{0}$$

$$\mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{X} \hat{\boldsymbol{\beta}} = \mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{y}$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{y} \quad (16)$$

Apabila disubstitusikan nilai  $\hat{\boldsymbol{\beta}}$  dari persamaan (16) ke persamaan (14) maka akan diperoleh dugaan untuk kurva  $\hat{g}$  seperti pada persamaan (17):

$$\begin{aligned}
\hat{g}(t) &= \mathbf{D}(\mathbf{y} - \mathbf{X} (\mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{y}) \\
&= \mathbf{D} \mathbf{y} - \mathbf{D} \mathbf{X} (\mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{y}
\end{aligned} \quad (17)$$

Dengan demikian diperoleh estimasi untuk kurva regresi semiparametrik deret Fourier seperti pada persamaan (18):

$$\begin{aligned}
\hat{y} &= \mathbf{X} \hat{\boldsymbol{\beta}} + \hat{g}(t) \\
&= (\mathbf{X} (\mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) + \\
&\quad \mathbf{D} - \mathbf{D} \mathbf{X} (\mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D}) \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{I} - \mathbf{D})^T (\mathbf{I} - \mathbf{D})) \mathbf{y} \\
&= \mathbf{A}(K) \mathbf{y}
\end{aligned} \quad (18)$$

## 2.8 Pemilihan K Optimal dengan GCV

Menurut Prahutama, Suparti dan Utami (2018) penentuan K optimal akan menghasilkan nilai  $R^2$  yang tinggi. Untuk memperoleh K optimal dapat dilihat dari nilai GCV yang paling minimum. Menurut Wu dan Zhang (2006) nilai GCV dapat dirumuskan seperti pada persamaan (31):

$$GCV(K) = \frac{MSE(K)}{(n^{-1} \text{trace}(\mathbf{I} - \mathbf{A}(K)))^2} \quad (19)$$

dengan:

$$MSE(K) = n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

K merupakan parameter osilasi,  $\mathbf{I}$  merupakan matriks identitas berukuran  $n \times n$  dan  $\mathbf{A}(K)$  merupakan matriks hat yang memuat parameter osilasi K.

## 2.9 Evaluasi Kinerja Model dan Peramalan

### 2.9.1 Koefisien Determinasi ( $R^2$ )

Adapun nilai dari koefisien determinasi adalah seperti pada persamaan (20):

$$R^2 = \frac{JKR}{JKT} = \frac{JKR}{JKR + JKS} \quad (20)$$

dengan:

$$JKR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

$$JKS = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Nilai  $R^2$  terletak diantara 0 dan 1. Jika  $R^2 = 1$ , maka model semakin baik menjelaskan variasi dalam  $y$ . Sebaliknya, jika  $R^2 = 0$ , berarti model tidak menjelaskan sedikitpun variasi dalam  $y$ . Kecocokan model lebih baik jika  $R^2$  semakin dekat dengan satu (Gujarati, 2004).

### 2.9.2 Mean Absolute Percentage Error (MAPE)

MAPE merupakan kuantitas error yang membandingkan data aktual dengan data prediksi dari hasil model regresi. Perhitungan MAPE sebagai berikut:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (21)$$

Menurut Lewis (1982) dalam Chen, *et al.* (2007) bahwa nilai MAPE yang dihasilkan mempunyai interpretasi sebagai berikut:

1. MAPE < 10% : peramalan sangat akurat
2. 10% ≤ MAPE < 20% : peramalan tersebut baik
3. 20% ≤ MAPE < 50% : peramalan masih dalam kewajaran
4. MAPE ≥ 50% : peramalan tidak akurat

## 2.10 Indeks Harga Saham Gabungan (IHSG)

Menurut Samsul (2006), Indeks Harga Saham Gabungan (*Composite Stock Price Index* = CSPI) merupakan indeks gabungan dari seluruh jenis saham yang tercatat di bursa efek. Indeks harga saham adalah suatu indikator yang menunjukkan pergerakan harga saham. Pasar yang sedang aktif, ditunjukkan dengan indeks harga saham yang mengalami kenaikan. Saat keadaan pasar stabil ditunjukkan dengan indeks harga saham yang tetap. Saat pasar sedang lesu ditunjukkan dengan indeks harga saham yang mengalami penurunan.

## 2.11 Indeks Dow Jones

Indeks *Dow Jones* merupakan salah satu indeks saham tertua di Amerika Serikat dan merupakan representasi dari kinerja industri terpenting di Amerika Serikat. Perusahaan yang tercatat di indeks *Dow Jones* pada umumnya merupakan perusahaan multinasional. Kegiatan operasi mereka tersebar ke seluruh dunia. Sebagai salah satu kekuatan ekonomi terbesar, pengaruh Amerika Serikat sangat besar bagi negara-negara lain termasuk Indonesia. Indeks *Dow Jones* yang bergerak naik, menandakan kinerja perekonomian Amerika Serikat secara umum berada pada posisi yang baik. Menurut Sunariyah (2006) dalam Wicaksono (2017) dengan adanya kondisi perekonomian yang baik, kegiatan perekonomian Indonesia akan digerakkan melalui kegiatan ekspor maupun melalui pasar

modal. Sehingga aliran modal yang masuk melalui pasar modal tentu akan memiliki pengaruh terhadap perubahan Indeks Harga Saham Gabungan (IHSG).

## 2.12 BI Rate

BI Rate adalah suku bunga kebijakan yang mencerminkan sikap atau *stance* kebijakan moneter yang ditetapkan oleh Bank Indonesia dan diumumkan kepada publik. Suku bunga yang ditetapkan oleh Bank Indonesia adalah sebagai patokan bagi suku bunga pinjaman maupun simpanan bagi bank atau lembaga keuangan di seluruh Indonesia. Suku bunga merupakan salah satu variabel yang dapat mempengaruhi harga saham karena perubahan tingkat suku bunga selanjutnya akan mempengaruhi keinginan seseorang untuk melakukan suatu investasi.

## 3. METODE PENELITIAN

### 3.1 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yaitu IHSG, Indeks *Dow Jones* dan BI Rate. Periode data yang digunakan yaitu bulanan terhitung dari bulan Juli 2010 sampai dengan Juni 2019 sebanyak 108 data. Data yang digunakan meliputi data *in sample* dan *out sample*. Data *in sample* terhitung mulai bulan Juli 2010 sampai dengan Juli 2018 sebanyak 97 data. Sedangkan data *out sample* terhitung mulai bulan Agustus 2018 sampai dengan Juni 2019 sebanyak 11 data. Data IHSG dan indeks *Dow Jones* diperoleh melalui situs <https://finance.yahoo.com/>. Sedangkan data BI Rate diperoleh dari website Badan Pusat Statistika [www.bps.go.id](http://www.bps.go.id) dan website Bank Indonesia [www.bi.go.id](http://www.bi.go.id). Variabel yang digunakan pada penelitian ini terdiri dari variabel respon dan prediktor, variabel-variabel tersebut sebagai berikut:

1. Indeks Harga Saham Gabungan sebagai variabel respon ( $y$ )
2. Indeks *Dow Jones* sebagai variabel prediktor parametrik ( $x$ )
3. BI Rate sebagai variabel prediktor nonparametrik ( $t$ )

### 3.1 Langkah-langkah Analisis Data

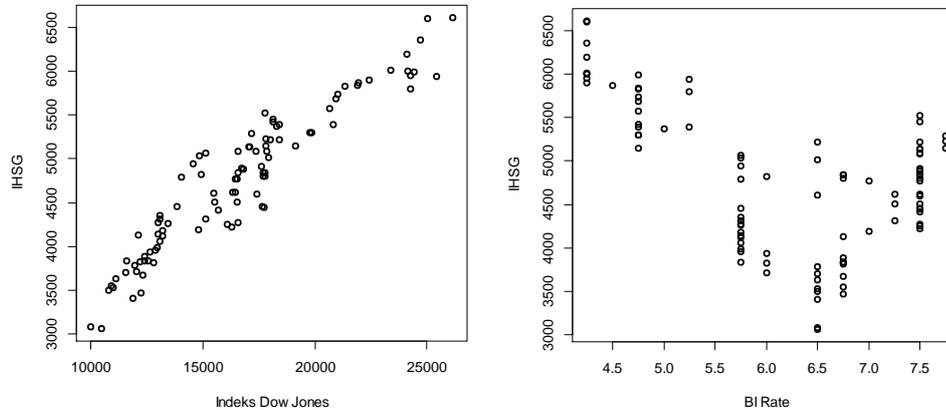
Software yang digunakan pada penelitian ini adalah R 3.5.1 dan *Microsoft Excel* 2013. Berikut langkah-langkah yang dilakukan:

1. Menghimpun data observasi yaitu IHSG, indeks *Dow Jones* dan BI Rate mulai bulan Juli 2010 sampai dengan Juni 2019.
2. Menentukan data *in sample* dan *out sample*.
3. Membuat *scatterplot* antara variabel respon dengan masing-masing variabel prediktor.
4. Menentukan variabel komponen parametrik dan nonparametrik.
5. Menentukan nilai parameter osilasi ( $K$ ) optimal berdasarkan kriteria GCV minimum.
6. Estimasi masing-masing parameter untuk komponen parametrik dan nonparametrik pada regresi semiparametrik dengan pendekatan deret fourier menggunakan metode *Ordinary Least Square*.
7. Memodelkan data dengan regresi semiparametrik deret fourier terbaik berdasarkan  $K$  optimal.
8. Menghitung nilai koefisien determinasi ( $R^2$ ) dan *Mean Absolute Percentage Error* (MAPE).
9. Menyimpulkan hasil penelitian.

## 4. HASIL DAN PEMBAHASAN

### 4.1 Plot antara IHSG dengan FaktorFaktor yang Mempengaruhinya

Berikut adalah scatterpot antara variabel respon (y) dengan masing-masing variabel prediktor (x dan t):



Berdasarkan *scatterplot* di atas data IHSG dengan indeks *Dow Jones* mendekati bentuk sebaran titik yang linier sehingga variabel indeks *Dow Jones* dikategorikan ke dalam komponen parametrik. Sedangkan *scatterplot* IHSG dan *BI Rate* membentuk pola acak atau tidak memiliki pola tertentu dan cenderung berulang, sehingga variabel *BI Rate* dikategorikan dalam komponen nonparametrik. Karena menggunakan dua faktor dengan komponen yang berbeda, maka untuk mendapatkan model terbaik metode yang tepat digunakan untuk analisis data tersebut adalah regresi semiparametrik. Metode regresi semiparametrik yang digunakan yaitu dengan pendekatan deret Fourier karena pada komponen nonparametrik pola data yang dihasilkan cenderung acak dan berulang.

### 4.2 Uji Signifikansi Model Regresi untuk Variabel Prediktor Komponen Parametrik

Berdasarkan output program R diperoleh nilai  $F_{hitung} = 816$  lebih besar dari  $F_{0,05; 1; 95} = 3,94$  dan  $p\text{-value} = 0,000$  lebih kecil dari  $\alpha = 0,05$ , sehingga dinyatakan bahwa  $H_0$  ditolak. Jadi, pada taraf signifikansi 5% dapat disimpulkan bahwa model sesuai yang berarti ada hubungan linier antara variabel respon IHSG dengan variabel prediktor indeks *Dow Jones*.

### 4.3 Model Regresi Semiparametrik dengan Pendekatan Deret Fourier

Untuk model regresi semiparametrik dengan pendekatan deret Fourier adalah sebagai berikut:

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \frac{1}{2} \sum_{l=1}^q \alpha_{0l} + \sum_{l=1}^q \gamma_l t_{il} + \sum_{l=1}^q \sum_{k=1}^K \alpha_{lk} \cos kt_{il} \quad (22)$$

Dalam penelitian ini variabel prediktor yang digunakan terdiri dari 1 komponen parametrik dan 1 komponen nonparametrik, maka model regresi semiparametrik dengan pendekatan deret Fourier dapat ditulis sebagai berikut:

$$y_i = \beta_0 + \beta_1 x_{i1} + \left( \frac{1}{2} \alpha_{01} + \gamma_1 t_{i1} + \sum_{k=1}^K \alpha_{1k} \cos kt_{i1} \right)$$

$$y_i = \omega + \beta_1 x_{i1} + \left( \frac{1}{2} \alpha_{01} + \gamma_1 t_{i1} + (\alpha_{11} \cos t_{i1} + \alpha_{12} \cos 2 t_{i1} + \dots + \alpha_{1k} \cos K t_{i1}) \right) \quad (23)$$

dengan  $\omega = \beta_0 + \phi$ ;  $\phi = \frac{1}{2} \alpha_{01}$

### 4.3.1 Penentuan K Optimal

Pemilihan K Optimal dilakukan dengan cara *trial* dan *error* hingga diperoleh nilai *Generalized Cross Validation* (GCV) minimum. Nilai K merupakan jumlah dari osilasi gelombang cosinus pada model. Nilai K yang semakin besar akan mengakibatkan model semakin kompleks dan osilasi dari kurva estimasi akan semakin rapat serta mengikuti pola data aktual.

**Tabel 2.** Nilai GCV setiap K yang dicobakan

No	K	GCV
1.	1	70608646
2.	2	39038712
3.	3	24943662
4.	4	16820196
5.	5	11918753
6.	6	9001069
7.	7	6682337
8.	8	5279168
9.	9	4332050
10.	10	4333896
11.	11	3639146
12.	12	3635366
13.	13	2826122
14.	14	2826122
15.	15	2826122

**Tabel 2** menunjukkan nilai GCV yang dihasilkan pada masing-masing nilai K yang dicobakan. Dari percobaan nilai K dari K = 1 sampai dengan K = 15 diperoleh nilai GCV minimum terdapat pada nilai K = 13 yaitu sebesar 2826122. Apabila nilai K dicobakan sampai dengan 100 akan diperoleh nilai GCV yang sama dengan nilai GCV pada K = 13 yaitu sebesar 2826122. Sehingga dapat disimpulkan bahwa nilai K optimal untuk penelitian ini adalah K = 13.

### 4.3.2 Estimasi Parameter pada Regresi Semiparametrik dengan Pendekatan Deret Fourier dengan K Optimal

Setelah diperoleh nilai K optimal selanjutnya dilakukan estimasi parameter untuk memperoleh model regresi semiparametrik dengan pendekatan deret Fourier. Nilai K optimal yang diperoleh adalah K = 13 maka banyaknya parameter yang harus ditaksir sebanyak 17 parameter dengan 2 parameter berasal dari komponen parametrik dan 15 parameter dari komponen nonparametrik. Berikut estimasi untuk masing-masing koefisien regresi semiparametrik deret Fourier untuk model Indeks Harga Saham Gabungan:

**Tabel 3.** Estimasi Parameter Koefisien Regresi

No	Parameter	Estimasi
1	$\beta_0$	$4,39 \times 10^{-19}$
2	$\beta_1$	0,209342
3	$\emptyset$	435388,914
4	$\gamma_1$	298,243
5	$\alpha_{11}$	-606408,745
6	$\alpha_{12}$	-26710,884
7	$\alpha_{13}$	629645,839
8	$\alpha_{14}$	-827858,818
9	$\alpha_{15}$	488413,7676
10	$\alpha_{16}$	206109,599
11	$\alpha_{17}$	-892268,15
12	$\alpha_{18}$	1246645,216
13	$\alpha_{19}$	-1177305,435
14	$\alpha_{110}$	824428,263
15	$\alpha_{111}$	-427029,281
16	$\alpha_{112}$	151991,909
17	$\alpha_{113}$	-29762,836

Sehingga estimasi untuk model regresi semiparametrik dengan pendekatan deret Fourier untuk Indeks Harga Saham Gabungan dapat ditulis:

$$\hat{y}_i = 4,39 \times 10^{-19} + 0,209342 x_1 + 435388,914 + 298,243 t_1 - 606408,745 \cos t_1 - 26710,884 \cos 2 t_1 + 629645,839 \cos 3 t_1 - 827858,818 \cos 4 t_1 + 488413,7676 \cos 5 t_1 + 206109,599 \cos 6 t_1 - 892268,15 \cos 7 t_1 + 1246645,216 \cos 8 t_1 - 1177305,435 \cos 9 t_1 + 824428,263 \cos 10 t_1 - 427029,281 \cos 11 t_1 + 151991,909 \cos 12 t_1 - 29762,836 \cos 13 t_1$$

#### 4.4 Ketepatan Kinerja Model dan Peramalan

##### 4.4.1 Koefisien Determinasi ( $R^2$ )

$R^2$  data *in sample*:

$$R^2 = \frac{JKR}{JKT} = \frac{JKR}{JKR + JKS} = \frac{58659362,548}{58659362,548 + 4923862,346} = 0,9226$$

Berdasarkan nilai perhitungan koefisien determinasi ( $R^2$ ) tersebut, dapat dijelaskan bahwa variabel independen indeks *Dow Jones* ( $x$ ) dan *BI Rate* ( $t$ ) mempengaruhi variabel dependen *IHSG* ( $y$ ) pada data *in sample* sebesar 92,26% dan sisanya 7,74% dipengaruhi oleh faktor lain.

#### 4.4.2 Mean Absolute Percentage Error (MAPE)

MAPE merupakan kuantitas error yang membandingkan data aktual dengan data prediksi dari hasil model regresi semiparametrik dengan pendekatan deret Fourier. Perhitungan MAPE sebagai berikut:

MAPE data *in sample*:

$$MAPE = \frac{1}{97} \times \left( \begin{array}{l} \left| \frac{3069,28 - 3383,272}{3069,28} \right| + \left| \frac{3081,88 - 3288,813}{3081,88} \right| \\ + \left| \frac{3501,30 - 3450,703}{3501,30} \right| + \left| \frac{3635,32 - 3519,878}{3635,32} \right| + \\ \left| \frac{3531,21 - 3496,333}{3531,21} \right| + \dots + \left| \frac{5936,44 - 6276,531}{5936,44} \right| \end{array} \right) \times 100\%$$

$$= 3,8154\%$$

MAPE data *out sample*:

$$MAPE = \frac{1}{11} \times \left( \begin{array}{l} \left| \frac{6018,46 - 5208,858}{6018,46} \right| + \left| \frac{5796,55 - 7075,76}{5796,55} \right| \\ + \left| \frac{5831,65 - 6794,708}{5831,65} \right| + \left| \frac{6056,12 - 6709,221}{6056,12} \right| + \\ \left| \frac{6194,50 - 6246,365}{6194,50} \right| + \dots + \left| \frac{6358,63 - 6931,437}{6358,63} \right| \end{array} \right) \times 100\%$$

$$= 8,4782\%$$

Berdasarkan perhitungan MAPE untuk data *in sample* diperoleh nilai sebesar 3,8154% dan untuk data *out sample* diperoleh nilai sebesar 8,4782% yang menunjukkan bahwa nilai MAPE < 10% yang artinya model regresi semiparametrik dengan pendekatan deret fourier yang dihasilkan memiliki kinerja yang sangat akurat.

#### 4.5 Nilai Koefisien Determinasi (R<sup>2</sup>) dan Mean Square Error (MSE) pada Setiap K

**Tabel 4.** Nilai R<sup>2</sup> dan MSE untuk setiap K

Nilai K	R <sup>2</sup>	MSE
K = 1	0,8969647	67539,36
K = 2	0,8987252	66385,31
K = 3	0,8988919	66276,07
K = 4	0,9018208	64356,16
K = 5	0,9053081	62070,24
K = 6	0,9065972	61225,25
K = 7	0,9122395	57526,76
K = 8	0,9144044	56107,64
K = 9	0,9150108	55710,18
K = 10	0,9149744	55734,02

K = 11	0,9150335	55695,28
K = 12	0,9151219	55637,44
K = 13	0,9225604	50761,47

Berdasarkan **Tabel 4** terlihat bahwa untuk nilai K=1 sebenarnya sudah menghasilkan nilai R<sup>2</sup> yang cukup tinggi yaitu sebesar 89,69647% dan setiap peningkatan dari nilai K nilai R<sup>2</sup> juga mengalami peningkatan yang artinya nilai K berbanding lurus dengan nilai R<sup>2</sup>. Sedangkan setiap peningkatan dari nilai K nilai MSE mengalami penurunan yang artinya nilai K berbanding terbalik dengan nilai MSE. Apabila nilai K yang dipilih untuk pemodelan yang parsimoni (sederhana) maka nilai K yang dipilih adalah K=1 namun dalam penelitian ini digunakan nilai K=13 karena mencari nilai K paling optimal berdasarkan kriteria nilai GCV minimum untuk pemodelan regresi semiparametrik dengan pendekatan deret Fourier. Apabila dicobakan nilai K sampai dengan K=100 nilai R<sup>2</sup> dan MSE yang dihasilkan bernilai sama seperti pada hasil K = 13. Oleh karena itu nilai K optimal yang dipilih untuk pemodelan adalah K = 13. Semakin besar nilai K maka bentuk kurva yang dihasilkan semakin mulus atau mendekati nilai aktual dari data.

## 5. KESIMPULAN

Berdasarkan hasil dan pembahasan, dapat disimpulkan bahwa:

1. Persamaan regresi semiparametrik dengan pendekatan deret Fourier yang diperoleh untuk menduga IHSG dengan nilai K optimal 13 adalah sebagai berikut:

$$\begin{aligned} \hat{y}_i = & 4,39 \times 10^{-19} + 0,209342 x_1 + 435388,914 + 298,243 t_1 - 606408,745 \\ & \cos t_1 - 26710,884 \cos 2 t_1 + 629645,839 \cos 3 t_1 - 827858,818 \cos 4 t_1 \\ & + 488413,7676 \cos 5 t_1 + 206109,599 \cos 6 t_1 - 892268,15 \cos 7 t_1 \\ & + 1246645,216 \cos 8 t_1 - 1177305,435 \cos 9 t_1 + 824428,263 \cos 10 t_1 \\ & - 427029,281 \cos 11 t_1 + 151991,909 \cos 12 t_1 - 29762,836 \cos 13 t_1 \end{aligned}$$

2. Model regresi semiparametrik dengan pendekatan deret Fourier yang terbentuk dengan nilai K optimal 13 menghasilkan nilai GCV minimum sebesar 2826122 dengan evaluasi ketepatan kinerja model pada data *in sample* diperoleh nilai R<sup>2</sup> sebesar 92,26% dan MAPE sebesar 3,8154%. Pada data *out sample* diperoleh nilai MAPE sebesar 8,4782%. Artinya model yang dihasilkan mempunyai kinerja yang sangat bagus baik untuk data *in sample* maupun data *out sample*.

## DAFTAR PUSTAKA

- Asrini, L.J. dan Budiantara, I.N. 2014. *Fourier Series Semiparametric Regression Models (Case Study: The Production of Low Land Rice Irrigation in Central Java)*. ARPN Journal of Engineering and Applied Sciences Vol. 9: Hal. 1501-1506.
- Bilodeau, M. 1992. *Fourier Smoother and Additive Models*. The Canadian Journal of Statistics Vol. 3: Hal. 257 – 259.
- Budiantara, I.N. 2005. *Model Keluarga Spline Polinomial Truncated Dalam Regresi Semiparametrik*. Makalah Seminar Nasional Matematika, Jurusan Matematika Universitas Diponegoro, Semarang.

- Budiantara, I.N. 2014. *Pemodelan Regresi Nonparametrik Dan Semiparametrik Spline (Konsep, Metode Dan Aplikasinya)*. Prosiding Seminar Nasional Matematika. Jurusan Matematika Universitas Udayana, Bali. ISSN: 2406-9868 Hal 1-16.
- Chen, R.J.C., Bloomfield, P. dan Cabbage, F.W. 2007. *Comparing Forecasting Models in Tourism*. Journal of Hospitality and Tourism Research 2007. DOI: 10.1177/1096348007309566
- Eubank, R.L. 1999. *Nonparametric Regression and Spline Smoothing*. New York: Marcel Dekker.
- Gujarati, D.N. 2004. *Basic Econometrics, Fourth Edition*. New York : McGraw-Hill.
- Hardle, W. 1994. *Applied Nonparametric Regression*. Berlin: Humboldt University.
- Makridarkis, S., Wheelwright, S.C. dan Hyndman R.J. 1998. *Forecasting Methods and Applications, Third Edition*. New York: John Wiley and Sons.
- Montgomery, D.C. dan Runger, G.C. 2003. *Applied Statistics and Probability for Engineers, Third Edition*. United States of America: John Wiley and Sons.
- Montgomery, D.C., Peck, E.A. dan Vining G.G. 2012. *Introduction to Linear Regression Analysis, Fifth Edition*. United States of America: John Wiley and Sons.
- Prahutama, A. 2013. *Model Regresi Nonparametrik dengan Pendekatan Deret Fourier Pada Kasus Tingkat Pengangguran Terbuka di Jawa Timur*. Prosiding Seminar Nasional Statistika Universitas Diponegoro, Semarang. ISBN: 978-602-14387-0-1.
- Prahutama, A., Suparti dan Utami, T.W. 2018. *Modelling Fourier Regression for Time Series Data – A Case Study: Modelling Inflation in Foods Sector in Indonesia*. IOP Conf. Series: Journal of Physics: Conf. Series 974 (2018) 012067. Hal. 1-9. DOI: 10.1088/1742-6596/974/1/012067.
- Samsul, M. 2006. *Pasar Modal dan Manajemen Portofolio*. Surabaya: Erlangga.
- Sunardi, N., Ula, L. N. R. 2017. *Pengaruh BI Rate, Inflasi dan Kurs Terhadap Indeks Harga Saham Gabungan (IHSG)*. Jurnal Sekuritas (Saham, Ekonomi, Keuangan dan Investasi) Vol. 1, No. 2: Hal. 27-41, Desember 2017.
- Suparti, Santoso, R., Prahutama, A., Yasin, H. dan Devi, A.R. 2018. *Analisis Data Inflasi Indonesia Menggunakan Metode Fourier dan Wavelet Multiscale Autoregressive*. Prosiding Seminar Nasional VARIANSI 2018. Hal. 185-196.
- Wicaksono, I.S. dan Yasa, G.W. 2017. *Pengaruh FED Rate, Indeks Dow Jones, Nikkei 225, Hans Seng Terhadap Indeks Harga Saham Gabungan*. E-Jurnal Akuntansi Universitas Udayana Vol. 18, No. 1, Januari 2017. Hal. 358-385.
- Wu, H., dan Zhang, J.T. 2006. *Nonparametric Regression Methods for Longitudinal Data Analysis*. John Wiley & Sons, Inc., Canada.