ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 12, Nomor 1, Tahun 2023, Halaman 30 - 41
Online di: https://ejournal3.undip.ac.id/index.php/gaussian/

ANALISIS INDEKS HARGA SAHAM GABUNGAN DAN FAKTOR PENGARUHNYA MENGGUNAKAN PEMODELAN REGRESI SEMIPARAMETRIK KERNEL DILENGKAPI GUI-R

Arnisa Melani Kahar[1*], Suparti[2], Arief Rachman Hakim[3]
[1,2,3] Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro
***e-mail** : arnisamk@gmail.com

DOI: 10.14710/j.gauss.12.1.30-41

Article Info:
Received: 2022-09-15
Accepted: 2022-11-20
Available Online: 2023-05-4

Keywords:
IDX; Money Supply; Inflation; Kernel Semiparametric Regression; MSE; R^2; MAPE; GUI.

Abstract: Composite Stock Price Index (IDX) shows the movement of stock prices used by investors to determine their investment strategy. IDX movement is influenced by macroeconomic factors such as money supply and inflation, so regression analysis is used to determine the relationship between the variables. Based on the scatterplot, money supply is known as a parametric predictor variable as it has a linier line patterned scatterplot and inflation is a nonparametric predictor variable as it has a random patterned scatterplot, so semiparametric regression modelling is used for the analysis. Kernel regression was chosen to analyze the nonparametric component based on the random patterned scatterplot of inflation. This study aims to obtain the results of semiparametric kernel regression modelling analysis and to create a GUI to be applied to the analysis as a development of previous similar studies that still done based on CLI. This study uses monthly data from January 2013 to December 2020 with the proportion of in sample and out sample data distribution 87,5%:12,5%. Based on the smallest MSE value as the best model criteria, semiparametric regression model with triangle kernel function is the best model obtained with optimal bandwidth=3.24, R^2 $(97.15\%) > 67\%$ which means the model is strong and MAPE $(6.79\%) < 10\%$ which means that the forecasting results are very accurate. GUI has been created according to the needs of the modelling analysis implementation.

1. PENDAHULUAN

Pasar modal merupakan salah satu instrumen yang penting bagi perkembangan perekonomian pada suatu negara. Dasar pengukuran kinerja saham sebagai salah satu instrumen investasi pada pasar modal dapat dilihat melalui fitur indeksasi saham yang diberi nama Indeks Harga Saham Gabungan atau IHSG. Melalui IHSG, investor dapat melihat kondisi pasar, apakah sedang mengalami peningkatan atau penurunan, untuk menentukan strategi investasinya. Blanchard (2006) menyatakan secara umum bahwa terdapat beberapa faktor yang memengaruhi pergerakan IHSG, salah satunya adalah faktor ekonomi makro seperti jumlah uang beredar dan inflasi. Hubungan antara IHSG, jumlah uang beredar, dan inflasi dapat diketahui dengan melakukan analisis regresi. Terdapat tiga model regresi yang dapat digunakan, yaitu model regresi parametrik, model regresi nonparametrik, dan model regresi semiparametrik yang merupakan gabungan dari kedua model regresi lainnya. Berdasarkan studi awal yaitu membuat *scatterplot* dari masing-masing data faktor pengaruh terhadap data IHSG, diketahui kedua faktor pengaruh memiliki pola *scatterplot* yang berbeda, yaitu berpola mengikuti garis linier (parametrik) dan berpola acak (nonparametrik), sehingga digunakan metode regresi semiparametrik untuk mendapatkan model terbaik. Adapun dilihat dari plot data komponen nonparametrik yang acak, maka model nonparametrik yang akan digunakan adalah model kernel dengan fungsi kernel yang dapat digunakan yaitu fungsi kernel gaussian, epanechnikov, segitiga, dan *uniform*. Beberapa penelitian terkait pemodelan regresi semiparametrik kernel yang telah dilakukan sebelumnya antara lain adalah Nanda *et al.* (2016), Fitriani *et al.* (2015), dan Perdina (2012) yang mana ketiganya menerapkan pemodelan regresi semiparametrik yang dilakukan masih hanya menggunakan satu hingga dua fungsi kernel tertentu dan analisis dilakukan dengan program R yang berbentuk *Command Line Interface.* Seiring dengan perkembangan teknologi masa kini, pelaksanaan analisis statistika menggunakan metode komputasi harus semakin dikembangkan sehingga dilakukanlah pengembangan dengan menerapkan teknologi berbasis *Graphical User Interface* atau GUI pada analisis pemodelan regresi semiparametrik kernel.

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah pada penelitian ini adalah menentukan hasil analisis pemodelan regresi semiparametrik kernel dengan data yang digunakan yaitu IHSG, jumlah uang beredar, dan inflasi. Selain itu, penelitian yang dilakukan membahas terkait pembuatan dan penerapan GUI pada analisis pemodelan regresi semiparametrik kernel menggunakan keempat fungsi kernel yang telah disebutkan. Tujuan yang ingin dicapai pada penelitian ini adalah didapatkannya hasil analisis pemodelan regresi semiparametrik kernel dan terciptanya GUI yang dapat diterapkan pada analisis pemodelan regresi semiparametrik kernel untuk memudahkan peneliti yang ingin melakukan penelitian menggunakan pemodelan serupa.

2. TINJAUAN PUSTAKA

Indeks Harga Saham Gabungan dan Faktor-Faktor Pengaruh

Indeks Harga Saham Gabungan atau IHSG adalah suatu nilai yang digunakan untuk mengukur kinerja gabungan seluruh saham yang tercatat di bursa efek (Sunariyah, 2003). Blanchard (2006) menyatakan bahwa terdapat banyak faktor ekonomi makro yang dapat memengaruhi pergerakan IHSG, dua di antaranya adalah inflasi dan jumlah uang beredar. Inflasi adalah proses kenaikan harga barang umum yang berlaku dalam perekonomian (Sukirno, 2002). Jumlah uang beredar berdasarkan Sukirno (2002) terdiri dari mata uang dalam peredaran dan uang giral (secara terbatas) ditambah uang kuasi (secara luas).

Regresi Parametrik

Menurut Hardle (1994), salah satu metode pendekatan regresi yaitu parametrik. Analisis regresi parametrik mengasumsikan bahwa bentuk fungsi atau kurva regresinya sudah diketahui. Pola hubungan antara variabel respon dengan variabel prediktor dapat diidentifikasi melalui *scatterplot*, apabila pada *scatterplot* terlihat pola yang cenderung linier maka digunakanlah model parametrik linier. Berdasarkan Gujarati (2003), model regresi linier sederhana dengan n observasi dan satu variabel prediktor dapat ditulis sebagai:

$$\mathrm{y_i} = \alpha + \beta \mathrm{x_i} + \varepsilon_i \qquad \text{dengan } \mathrm{i} = 1, 2, \dots, \mathrm{n} \tag{1}$$

jika ditulis dalam bentuk maktriks, persamaan (1) dapat dirumuskan sebagai berikut:

$$\mathbf{Y} = \mathbf{Xb} + \boldsymbol{\varepsilon} \tag{2}$$

Parameter model regresi linier dapat dihitung nilai taksirannya menggunakan metode kuadrat terkecil yang dirumuskan pada persamaan berikut:

$$\sum_{\mathrm{i}=1}^{\mathrm{n}} \varepsilon_i^2 = \sum_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{y_i} - \hat{\mathrm{y}}_\mathrm{i})^2 = \sum_{\mathrm{i}=1}^{\mathrm{n}} \left(\mathrm{y_i} - \hat{\alpha} - \hat{\beta}\mathrm{x_i}\right)^2 \tag{3}$$

Untuk mencari $\hat{\alpha}$ dan $\hat{\beta}$ menggunakan metode kuadrat terkecil adalah dengan mencari turunan pertama dari persamaan (3) dan disamadengankan nol, didapatkan hasil berikut:

$$\hat{\alpha} = \frac{\sum_{i=1}^{n} \mathrm{x_i}^2 \sum_{i=1}^{n} \mathrm{y_i} - \sum_{i=1}^{n} \mathrm{x_i} \sum_{i=1}^{n} \mathrm{x_i y_i}}{\mathrm{n}(\sum_{i=1}^{n} \mathrm{x_i}^2 - (\sum_{i=1}^{n} \mathrm{x_i})^2} \tag{4}$$

$$\hat{\beta} = \frac{\mathrm{n} \sum_{i=1}^{n} \mathrm{x_i y_i} - \sum_{i=1}^{n} \mathrm{x_i} \sum_{i=1}^{n} \mathrm{y_i}}{\mathrm{n}(\sum_{i=1}^{n} \mathrm{x_i}^2) - (\sum_{i=1}^{n} \mathrm{x_i})^2} \tag{5}$$

Hasil estimasi $\widehat{\alpha}$ dan $\widehat{\beta}$ dapat dikatakan memiliki jumlah kuadrat residual paling minimum jika telah memenuhi syarat $\frac{\partial^2 \sum_{i=1}^{n} \varepsilon_i{}^2}{\partial \widehat{\alpha}^2} > 0$ dan $\frac{\partial^2 \sum_{i=1}^{n} \varepsilon_i{}^2}{\partial \widehat{\beta}^2} > 0$ dengan pembuktian berikut:

$$\frac{\partial^2 \sum_{i=1}^{n} \varepsilon_i{}^2}{\partial \widehat{\alpha}^2} = \frac{\partial^2 \sum_{i=1}^{n} (y_i - \widehat{\alpha} - \widehat{\beta} x_i)^2}{\partial \widehat{\alpha}^2} = 2n > 0$$

$$\frac{\partial^2 \sum_{i=1}^{n} \varepsilon_i{}^2}{\partial \widehat{\beta}^2} = \frac{\partial^2 \sum_{i=1}^{n} (y_i - \widehat{\alpha} - \widehat{\beta} x_i)^2}{\partial \widehat{\beta}^2} = 2 \sum_{i=1}^{n} x_i{}^2 > 0$$

Jika ditulis secara sederhana dalam bentuk matriks, estimasi model regresi mengacu pada persamaan (2) adalah sebagai berikut:

$$\widehat{\mathbf{Y}} = \mathbf{X}\,\widehat{\mathbf{b}} \tag{6}$$

$$\widehat{\mathbf{b}} = (\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\,\mathbf{X}^{\mathbf{T}}\mathbf{Y} \tag{7}$$

Regresi Nonparametrik

Hardle (1994) menyatakan bahwa konsep regresi nonparametrik tidak menuntut terpenuhi banyak asumsi seperti regresi parametrik, salah satunya asumsi normalitas. Pendekatan regresi nonparametrik merupakan suatu metode yang dilakukan untuk memodelkan data yang tidak diketahui bentuk fungsinya. Model umum regresi nonparametrik dapat dituliskan sebagai berikut:

$$y_i = m(t_i) + \varepsilon_i \qquad \text{dengan } i = 1, 2, \ldots, n \tag{8}$$

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

$$\mathbf{Y} = \mathbf{m} + \boldsymbol{\varepsilon}$$

Fungsi Densitas Kernel

Ada beberapa teknik pendugaan $m(t_i)$ dalam regresi nonparametrik, salah satunya adalah metode kernel. Berdasarkan Wand dan Jones (1995), estimator densitas kernel dengan fungsi kernel K dan *bandwidth* h didefinisikan sebagai berikut:

$$\hat{f}(x; h) = (nh)^{-1} \sum_{i=1}^{n} K\{(\mathbf{x} - x_i)/h\} \tag{9}$$

Ketepatan estimator ini dapat disesuaikan dengan mengubah *bandwidth* h, jika h kecil maka akan dikatakan *undersmooth* sedangkan jika h besar maka akan dikatakan *oversmooth* (Ogden dalam Suparti *et al.*, 2018). Carmona (2004) menjelaskan bahwa $K(\cdot)$ dapat dikatakan fungsi kernel apabila memiliki sifat-sifat berikut:

1. $\int_{-\infty}^{\infty} K(x)\, dx = 1$
2. $K(-x) = K(x)$
3. $\int_{-\infty}^{\infty} x\, K(x)\, dx = 0$

Terdapat beberapa fungsi kernel yang biasa digunakan, di antaranya adalah:

1. Gaussian
   $$K(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \text{untuk } |x| < \infty$$
2. Epanechnikov
   $$K(x) = \begin{cases} \frac{3}{4}(1 - x^2), \text{untuk } |x| \le 1 \\ 0, \text{untuk x lainnya} \end{cases}$$
3. Segitiga
   $$K(x) = \begin{cases} 1 - |x|, \text{untuk } |x| \le 1 \\ 0, \text{untuk x lainnya} \end{cases}$$
4. Seragam (*Uniform*)
   $$K(x) = \begin{cases} \frac{1}{2}, \text{untuk } |x| \le 1 \\ 0, \text{untuk x lainnya} \end{cases}$$

Regresi Nonparametrik Kernel

Regresi nonparametrik kernel adalah teknik statistika nonparametrik untuk memperkirakan ekspektasi bersyarat dari variabel acak dengan tujuan untuk menemukan hubungan antara variabel acak T dan Y dengan menggunakan bobot fungsi kernel. Misalkan diberikan data (t_i, y_i) dengan $i = 1,2,\dots,n$ dan akan diestimasi sebuah fungsi regresi $m(t)$ yang akan memenuhi model $y_i = m\ t_i + \varepsilon_i$, nilai $m(t)$ ekuivalen dengan nilai harapan dari variabel respon jika variabel prediktor $T = t$ telah diketahui. Secara matematis dapat ditulis sebagai berikut:

$$m(t) = E(Y|T = t) = \int_{-\infty}^{\infty} yf(y|t)dy = \int_{-\infty}^{\infty} \frac{y\,f(t,y)}{f(t)} dy \tag{10}$$

pendugaan f(t,y) dan f(t) adalah sebagai berikut:

$$\hat{f}(t,h) = \frac{1}{nh_t h_y} \sum_{i=1}^{n} K_t\left(\frac{t-t_i}{h_t}\right) K_y\left(\frac{y-y_i}{h_y}\right) \tag{11}$$

$$\hat{f}(t) = \frac{1}{nh_t} \sum_{i=1}^{n} K_t\left(\frac{t-t_i}{h_t}\right) \tag{12}$$

dengan $K_t(.)$ dan $K_y(.)$ adalah fungsi kernel sedangkan h_t dan h_y adalah *bandwidth.* Sehingga $\widehat{m}(t)$ diperoleh dengan mensubstitusi persamaan (11) dan (12) ke dalam persamaan (10) menjadi:

$$\widehat{m}(t) = \int_{-\infty}^{\infty} \frac{y\frac{1}{nh_t h_y} \sum_{i=1}^{n} K_t\left(\frac{t-t_i}{h_t}\right) K_y\left(\frac{y-y_i}{h_y}\right)}{\frac{1}{nh_t} \sum_{i=1}^{n} K_t\left(\frac{t-t_i}{h_t}\right)} dy \tag{13}$$

misalkan $\eta = \frac{y-y_i}{h_y}$ maka diperoleh:

$$\frac{1}{h_y} \int_{-\infty}^{\infty} y\, K_y\left(\frac{y-y_i}{h_y}\right) dy = \int_{-\infty}^{\infty} (h_y\eta + y_i) K_y(\eta) d\eta = y_i \tag{14}$$

dengan mensubstitusi persamaan (14) ke dalam persamaan (13) dan untuk $h_t = h$, maka diperoleh estimator Nadaraya-Watson sebagai berikut:

$$\widehat{m}(t) = \frac{\sum_{i=1}^{n} K\left(\frac{t-t_i}{h}\right) y_i}{\sum_{i=1}^{n} K\left(\frac{t-t_i}{h}\right)}$$

$$\widehat{m}(t) = \frac{\sum_{i=1}^{n} K\left(\frac{t-t_i}{h}\right) y_i}{\sum_{k=1}^{n} K\left(\frac{t-t_k}{h}\right)} = \sum_{i=1}^{n} W_i(t) y_i \tag{15}$$

dengan $W_i(t) = \frac{K\left(\frac{t-t_i}{h}\right)}{\sum_{k=1}^{n} K\left(\frac{t-t_k}{h}\right)}$.

$W_i(t)$ merupakan barisan bobot positif dan memiliki karakteristik $\sum_{i=1}^{n} W_i(t) = 1$ sehingga estimator Nadaraya-Watson merupakan rata-rata terbobot dari y_i. Dengan mengganti indeks i dengan j, persamaan (15) dapat ditulis sebagai berikut:

$$\widehat{m}(t) = \sum_{j=1}^{n} W_j(t) y_j$$

Ketika estimasi $t = t_i$, maka $\widehat{m}(t_i)$ disebut $\hat{y}_i$ dan diperoleh:

$$\hat{y}_i = \sum_{j=1}^{n} W_j(t_i) y_j \tag{16}$$

dengan $W_j(t_i) = \frac{K\left(\frac{t_i-t_j}{h}\right)}{\sum_{k=1}^{n} K\left(\frac{t_i-t_k}{h}\right)}$.

Regresi Semiparametrik

Hardle *et al.* (2000) menjelaskan bahwa model regresi semiparametrik ini dapat digunakan pada kasus yang terdiri dari komponen parametrik yang diketahui polanya dan komponen nonparametrik yang tidak diketahui polanya. Bentuk model umum regresi semiparametrik dapat dituliskan sebagai berikut:

$$y_i = \alpha + \beta x_i + m(t_i) + \varepsilon_i \qquad \text{dengan } i = 1, 2, \dots, n \tag{17}$$

jika ditulis dalam bentuk maktriks, persamaan (17) dapat dirumuskan sebagai berikut:

$$\mathbf{Y} = \mathbf{Xb} + \mathbf{m} + \boldsymbol{\varepsilon} \tag{18}$$

Estimasi Model Regresi Semiparametrik Kernel

Untuk memperoleh model penduga semiparametrik kernel berdasarkan persamaan (17), fungsi m dapat diestimasi terlebih dahulu dengan estimator Nadaraya-Watson berikut:

$$\widehat{m}(t_i) = \frac{\sum_{j=1}^{n} K\left(\frac{t_i - t_j}{h}\right) y_j}{\sum_{k=1}^{n} K\left(\frac{t_i - t_k}{h}\right)}$$

sedangkan untuk mengestimasi parameter $\mathbf{b}$ dapat dilakukan permisalan dari persamaan (17) dengan menggunakan estimator Nadaraya-Watson untuk fungsi $m(t_i)$ sebagai berikut:

$$y_i = \alpha + \beta x_i + \widehat{m}(t_i) + \varepsilon_i$$

$$y_i - \widehat{m}(t_i) = \alpha + \beta x_i + \varepsilon_i$$

$$y_i^* = \alpha + \beta x_i + \varepsilon_i \tag{19}$$

Jika $y_i^* = y_i - \widehat{m}(t_i)$, maka nilai $\widehat{\alpha}$ dan $\widehat{\beta}$ dapat dihitung menggunakan persamaan (4) dan (5) kemudian mensubstitusi nilai y_i dengan y_i^* menjadi seperti berikut:

$$\widehat{\beta} = \frac{n\sum_{i=1}^{n} x_i y_i^* - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i^*}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$\widehat{\beta} = \frac{n\sum_{i=1}^{n} x_i\left(y_i - \widehat{m}(t_i)\right) - \sum_{i=1}^{n} x_i \sum_{i=1}^{n}\left(y_i - \widehat{m}(t_i)\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{20}$$

$$\widehat{\alpha} = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^* - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i^*}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$\widehat{\alpha} = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n}\left(y_i - \widehat{m}(t_i)\right) - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i\left(y_i - \widehat{m}(t_i)\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{21}$$

dalam bentuk matriks persamaan (19) dapat ditulis sebagai berikut:

$$\mathbf{Y}^* = \mathbf{X}\,\mathbf{b} + \boldsymbol{\varepsilon}$$

$$\hat{\mathbf{b}} = (\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\,\mathbf{X}^{\mathbf{T}}\mathbf{Y}^*$$

$$\hat{\mathbf{Y}}^* = \mathbf{X}\hat{\mathbf{b}}$$

sehingga untuk mengestimasikan model regresi semiparametrik kernel, dalam matriks dapat dituliskan sebagai berikut ini:

$$\hat{\mathbf{Y}} = \hat{\mathbf{Y}}^* + \widehat{\mathbf{m}} \tag{22}$$

***Mean Squared Error* (MSE)**

Penentuan model regresi semiparametrik kernel terbaik dapat dibandingkan melalui nilai rata-rata jumlah kuadrat residual atau MSE yang diperoleh masing-masing model regresi. Model dapat dikatakan sebagai model regresi terbaik apabila memiliki nilai MSE terkecil. Berdasarkan Kurniasih *et al.* (2013), nilai MSE dapat didefinisikan sebagai berikut:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{23}$$

Koefisien Determinasi (R^2)

Menurut Gujarati (2003), koefisien determinasi merupakan ukuran ikhtisar yang mengatakan seberapa baik garis regresi sampel mencocokkan data. Nilai R^2 dapat didefinisikan melalui persamaan berikut:

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{24}$$

Batas nilai koefisien determinasi adalah $0 \leq R^2 \leq 1$. Suatu R^2 bernilai sebesar 1 berarti suatu kecocokan sempurna, sedangkan R^2 yang bernilai nol berarti tidak ada hubungan antara variabel tak bebas dengan variabel yang menjelaskan (Gujarati, 2003). Berdasarkan

Chin (1998), kriteria nilai R^2 yang telah didapatkan dari perhitungan menggunakan persamaan (24) dapat dilihat pada Tabel 1.

Tabel 1. Tabel Kriteria Nilai R^2

Kriteria R^2	Interpretasi
$19\% < R^2 \leq 33\%$	Model lemah
$33\% < R^2 \leq 67\%$	Model moderat
$R^2 > 67\%$	Model kuat

Mean Absolute Percentage Error (MAPE)

Makridakis *et al.* (1997) menjelaskan bahwa dalam kebanyakan situasi peramalan, akurasi diperlukan sebagai kriteria utama untuk memilih metode peramalan. Terdapat beberapa cara yang dapat digunakan untuk mengukur tingkat akurasi peramalan, salah satunya adalah *Mean Absolute Percentage Error* atau MAPE yang didefinisikan sebagai berikut:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\left(\frac{y_i-\hat{y}_i}{y_i}\right)\times 100\%\right| \tag{25}$$

Berdasarkan Lewis (1982), nilai MAPE yang telah didapatkan dari perhitungan menggunakan persamaan (25) dapat diinterpretasikan menjadi empat kategori dan dapat dilihat pada Tabel 2.

Tabel 2. Tabel Interpretasi Nilai MAPE

MAPE (%)	Interpretasi
MAPE < 10%	Prediksi sangat akurat
10% ≤ MAPE < 20 %	Prediksi baik
20% ≤ MAPE < 50 %	Prediksi wajar
MAPE ≥ 50 %	Prediksi tidak akurat

Graphical User Interface (GUI)

Program R tidak hanya dapat digunakan untuk melakukan komputasi statistika berbasis *Command Line Interface* (CLI) saja, tetapi juga dapat membuat visualisasi hasil analisisnya dalam bentuk *Graphical User Interface* (GUI). Sama seperti komputasi lainnya yang menggunakan program R, untuk membuat *interface* ini diperlukan suatu paket R yang memang dirancang khusus untuk pembuatan GUI, salah satu di antaranya adalah *Shiny Package*. Pembuatan aplikasi GUI menggunakan paket *Shiny* memerlukan tiga komponen utama, yaitu UI (*User Interface*) untuk mengatur tampilan aplikasi, *Server* untuk mendefinisikan program yang ingin dibuat dengan menyesuaikan tampilan UI, serta fungsi *ShinyApp* untuk memanggil UI dan *Server* yang telah didefinisikan sebelumnya sehingga aplikasi yang telah dibuat dapat dijalankan.

3. METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder yang terdiri dari Indeks Harga Saham Gabungan (IHSG) sebagai variabel respon (y), jumlah uang beredar sebagai variabel prediktor komponen parametrik (x), dan inflasi sebagai variabel prediktor komponen nonparametrik (t). Data merupakan data bulanan mulai dari bulan Januari 2013 sampai dengan bulan Desember 2020 dan dibagi menjadi data *in sample* dan *out sample* dengan proporsi pembagian sebesar 87,5% : 12,5%. *Software* statistika yang digunakan dalam penelitian adalah RStudio, Minitab, dan Microsoft Excel 2013.

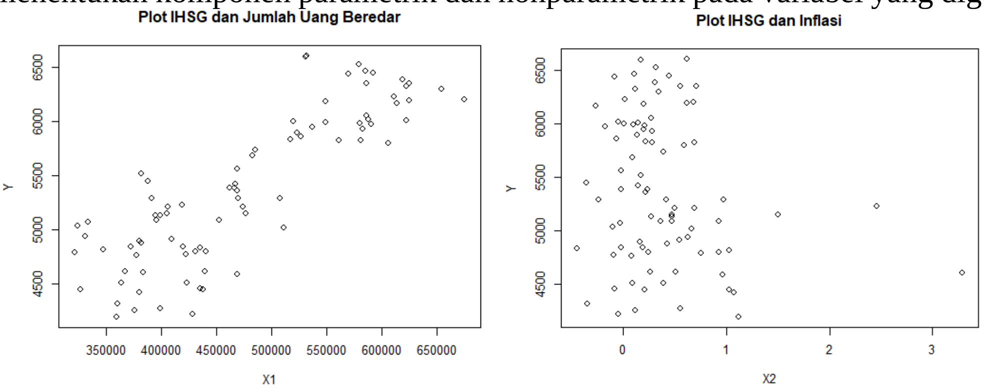
Tahapan penelitian yang dilakukan untuk menganalisis pengaruh variabel prediktor terhadap variabel respon menggunakan pemodelan regresi semiparametrik kernel sehingga didapatkan model regresi terbaik adalah sebagai berikut:

1. Membuat aplikasi GUI sesuai kebutuhan analisis
2. Menetapkan variabel penelitian dan mengumpulkan data variabel yang diperlukan

3. Menentukan data *in sample* dan *out sample*
4. Mendeskripsikan data dengan *scatterplot* menggunakan data *in sample*
5. Menentukan variabel yang menjadi komponen parametrik dan komponen nonparametrik sesuai *scatterplot* yang diperoleh
6. Menentukan fungsi kernel yang digunakan
7. Menentukan besar *bandwidth* h
8. Menghitung estimasi parameter **b** tiap *bandwidth* h
9. Menghitung nilai MSE tiap model regresi semiparametrik yang terbentuk
10. Menentukan model regresi terbaik berdasarkan nilai MSE minimum
11. Menghitung nilai koefisien determinasi data *in sample* sebagai ukuran kebaikan model
12. Menghitung nilai prediksi data *out sample* menggunakan model regresi terbaik
13. Menghitung nilai MAPE menggunakan nilai prediksi data *out sample* sebagai evaluasi kinerja model
14. Menyimpulkan hasil penelitian.

4. HASIL DAN PEMBAHASAN

Pemodelan menggunakan model regresi semiparametrik dapat dimulai dengan membuat *scatterplot* dari data *in sample* variabel respon dengan masing-masing variabel prediktornya untuk menentukan komponen parametrik dan nonparametrik pada variabel yang digunakan.



(a) (b)

Gambar 1. (a) *Scatterplot* IHSG dan Jumlah Uang Beredar, (b) *Scatterplot* IHSG dan Inflasi Januari 2013 – Desember 2019

Data yang digunakan sebagai variabel respon pada penelitian ini adalah data nilai IHSG, variabel prediktor komponen parametrik adalah jumlah uang beredar karena berdasarkan Gambar 1 (a) terlihat bahwa *scatterplot* membentuk suatu pola yang mengikuti garis linier sedangkan inflasi merupakan variabel prediktor komponen nonparametrik berdasarkan pada Gambar 1 (b) yang menunjukkan *scatterplot* yang cenderung berpola acak.

Berdasarkan *trial and error* yang dilakukan dengan program R menggunakan fungsi kernel gaussian, didapatkan model semiparametrik kernel gaussian dengan nilai MSE minimum yang hasilnya dapat dilihat pada Tabel 3.

Tabel 3. Lima Nilai *Bandwidth*, Parameter **b**, dan MSE dengan Fungsi Kernel Gaussian

No.	*Bandwidth* h	$\hat{\alpha}$	$\hat{\beta}$	MSE
1.	1,30	-2925,5805	0,00612273	131556,8502
2.	1,29	-2925,2573	0,00612183	131557,0796
3.	1,31	-2925,8922	0,00612361	131557,1946
4.	1,28	-2924,9221	0,00612089	131557,9305
5.	1,32	-2926,1927	0,00612447	131558,0685

Tabel 3 menunjukkan beberapa model yang telah diujikan nilai MSE-nya. Model regresi dapat dikatakan sebagai model terbaik apabila memiliki nilai MSE terkecil, sehingga dari

Tabel 3 dapat diketahui model regresi semiparametrik dengan fungsi kernel gaussian terbaik serta dapat ditulis dalam persamaan mengacu pada persamaan (17) sebagai berikut:

$$\hat{y}_i = -2925{,}5805 + 0{,}00612273x_i + \frac{\sum_{j=1}^{84} K\left(\frac{t_i - t_j}{1{,}30}\right) y_j}{\sum_{k=1}^{84} K\left(\frac{t_i - t_k}{1{,}30}\right)} \quad (26)$$

Berdasarkan *trial and error* yang dilakukan dengan program R menggunakan fungsi kernel epanechnikov, didapatkan model semiparametrik kernel epanechnikov dengan nilai MSE minimum yang hasilnya dapat dilihat pada Tabel 4.

Tabel 4. Lima Nilai *Bandwidth*, Parameter **b**, dan MSE dengan Fungsi Kernel Epanechnikov

No.	***Bandwidth* h**	$\hat{\boldsymbol{\alpha}}$	$\hat{\boldsymbol{\beta}}$	**MSE**
1.	3,18	-2924,0785	0,00613414	131632,7818
2.	3,17	-2923,7126	0,00613337	131633,3517
3.	3,19	-2924,4684	0,00613494	131633,5828
4.	3,16	-2923,3350	0,00613259	131633,9425
5.	3,20	-2924,8694	0,00613576	131635,4069

Tabel 4 menunjukkan beberapa model yang telah diujikan nilai MSE-nya. Model regresi dapat dikatakan sebagai model terbaik apabila memiliki nilai MSE terkecil, sehingga dari Tabel 4 dapat diketahui model regresi semiparametrik dengan fungsi kernel epanechnikov terbaik serta dapat ditulis dalam persamaan mengacu pada persamaan (17) sebagai berikut:

$$\hat{y}_i = -2924{,}0785 + 0{,}00613414x_i + \frac{\sum_{j=1}^{84} K\left(\frac{t_i - t_j}{3{,}18}\right) y_j}{\sum_{k=1}^{84} K\left(\frac{t_i - t_k}{3{,}18}\right)} \quad (27)$$

Berdasarkan *trial and error* yang dilakukan dengan program R menggunakan fungsi kernel segitiga, didapatkan model semiparametrik kernel segitiga dengan nilai MSE minimum yang hasilnya dapat dilihat pada Tabel 5.

Tabel 5. Lima Nilai *Bandwidth*, Parameter **b**, dan MSE dengan Fungsi Kernel Segitiga

No.	***Bandwidth* h**	$\hat{\boldsymbol{\alpha}}$	$\hat{\boldsymbol{\beta}}$	**MSE**
1.	3,24	-2920,1440	0,00611025	130908,7797
2.	3,25	-2920,4682	0,00611094	130908,8922
3.	3,23	-2919,8062	0,00610953	130909,2312
4.	3,26	-2920,7795	0,00611161	130909,5077
5.	3,22	-2919,4538	0,00610878	130910,3156

Tabel 5 menunjukkan beberapa model yang telah diujikan nilai MSE-nya. Model regresi dapat dikatakan sebagai model terbaik apabila memiliki nilai MSE terkecil, sehingga dari Tabel 5 dapat diketahui model regresi semiparametrik dengan fungsi kernel segitiga terbaik serta dapat ditulis dalam persamaan mengacu pada persamaan (17) sebagai berikut:

$$\hat{y}_i = -2920{,}1440 + 0{,}00611025x_i + \frac{\sum_{j=1}^{84} K\left(\frac{t_i - t_j}{3{,}24}\right) y_j}{\sum_{k=1}^{84} K\left(\frac{t_i - t_k}{3{,}24}\right)} \quad (28)$$

Berdasarkan *trial and error* yang dilakukan dengan program R menggunakan fungsi kernel *uniform*, didapatkan model semiparametrik kernel *uniform* dengan nilai MSE minimum yang hasilnya dapat dilihat pada Tabel 6.

Tabel 6. Lima Nilai *Bandwidth*, Parameter **b**, dan MSE dengan Fungsi Kernel *Uniform*

No.	***Bandwidth* h**	$\hat{\boldsymbol{\alpha}}$	$\hat{\boldsymbol{\beta}}$	**MSE**
1.	2,78	-2923,6542	0,00614431	131319,9415
2.	2,79	-2923,2017	0,00614357	131376,6320
3.	2,80	-2923,2017	0,00614357	131376,6320
4.	2,81	-2922,3952	0,00614225	131415,0489
5.	2,75	-2925,6652	0,00614771	131458,2095

Tabel 6 menunjukkan beberapa model yang telah diujikan nilai MSE-nya. Model regresi dapat dikatakan sebagai model terbaik apabila memiliki nilai MSE terkecil, sehingga dari

Tabel 6 dapat diketahui model regresi semiparametrik dengan fungsi kernel *uniform* terbaik serta dapat ditulis dalam persamaan mengacu pada persamaan (17) sebagai berikut:

$$\hat{y}_i = -2923{,}6542 + 0{,}00614431x_i + \frac{\sum_{j=1}^{84} K\left(\frac{t_i - t_j}{2{,}78}\right) y_j}{\sum_{k=1}^{84} K\left(\frac{t_i - t_k}{2{,}78}\right)} \tag{29}$$

Analisis pemodelan regresi semiparametrik kernel dilakukan dengan cara meminimumkan nilai MSE dari masing-masing model yang terbentuk. Dari model terbaik masing-masing regresi semiparametrik menggunakan keempat fungsi kernel, diperoleh nilai MSE seperti pada Tabel 7.

Tabel 7. Perbandingan Nilai MSE Model Regresi Semiparametrik Kernel

Fungsi Kernel	MSE
Gaussian	131556,8502
Epanechnikov	131632,7818
Segitiga	130908,7797
Uniform	131319,9415

Berdasarkan Tabel 7 terlihat bahwa model kernel segitiga memiliki nilai MSE terkecil di antara keempat model, yaitu sebesar 130908,7797. Hal tersebut menunjukkan bahwa model regresi semiparametrik dengan fungsi kernel segitiga merupakan model terbaik.

Setelah model regresi terbaik diperoleh, R^2 dihitung untuk mengetahui seberapa baik garis regresi sesuai dengan data aktualnya. Data yang digunakan merupakan data *in sample* dari IHSG. Perhitungan nilai R^2 untuk fungsi kernel segitiga sebagai model regresi terbaik mengacu pada persamaan (24) dapat dilihat sebagai berikut:

$$R^2 = \frac{(4452{,}88-5383{,}88)^2+(4435{,}75-5383{,}88)^2+\cdots+(6488{,}76-5383{,}88)^2}{(4453{,}70-5383{,}88)^2+(4795{,}79-5383{,}88)^2+\cdots+(6299{,}54-5383{,}88)^2} = 0{,}971507$$

Diperoleh nilai R^2 sebesar 0,971507 yang dapat diartikan bahwa variabel prediktor (x) dan (t) memengaruhi variabel respon (y) sebesar 97,15% dan sebesar 2,85% lainnya dipengaruhi oleh faktor lain serta model dikategorikan sebagai model kuat karena nilai R^2 > 67%.

Perhitungan MAPE dilakukan untuk mengetahui tingkat akurasi peramalan. Data yang digunakan merupakan data *out sample*. Perhitungan nilai MAPE untuk model regresi semiparametrik kernel segitiga sebagai model terbaik mengacu pada persamaan (25) dapat dilihat sebagai berikut ini:

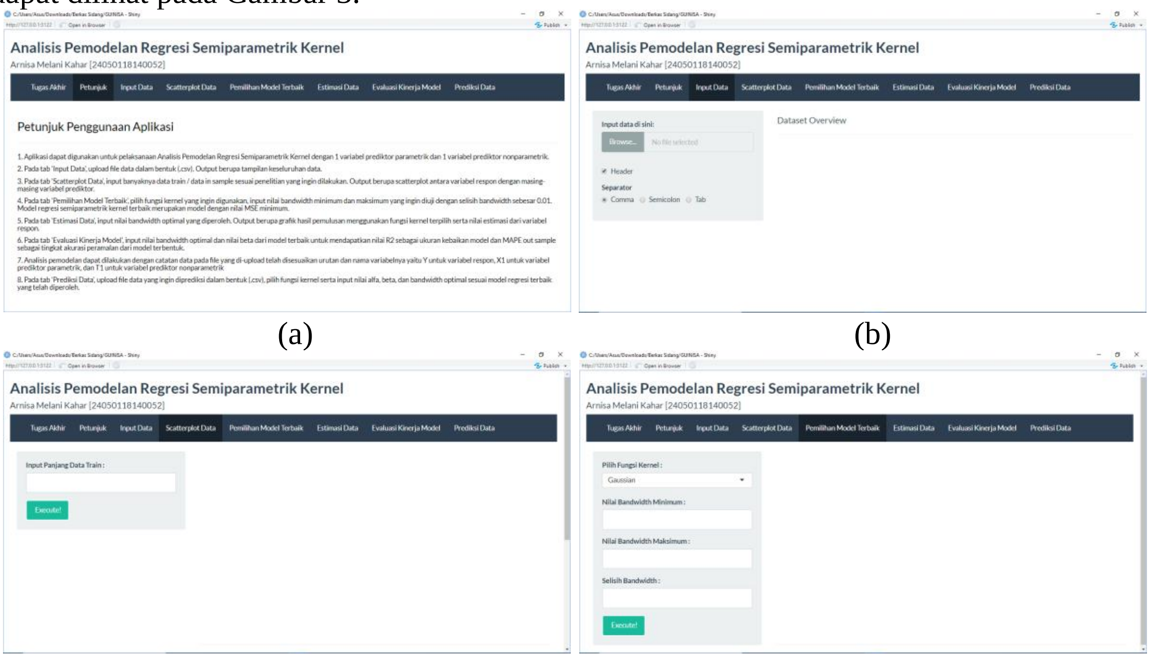
$$\text{MAPE} = \frac{1}{12}\left|\left(\frac{5940{,}05-5043{,}14}{5940{,}05} + \cdots + \frac{5979{,}07-5528{,}81}{5979{,}07}\right) \times 100\%\right| = 6{,}7912\%$$

Diperoleh nilai MAPE sebesar 6,79% yang menunjukkan bahwa nilai MAPE tersebut berada pada kategori nilai MAPE < 10% yang berarti hasil peramalan dari model regresi semiparametrik kernel terbaik yang diperoleh sangat akurat.

Pembuatan GUI menggunakan paket *shiny* memerlukan tiga komponen utama yaitu UI, *Server*, dan fungsi *ShinyApp*. Desain tampilan aplikasi GUI yang dibangun didefinisikan pada objek UI dan dibangun dengan menyesuaikan kebutuhan pelaksanaan analisis pemodelan regresi semiparametrik kernel. Pada bagian *Server*, logika kerja analisis yang dibutuhkan untuk pelaksanaan analisis didefinisikan sesuai dengan UI yang telah dibuat.

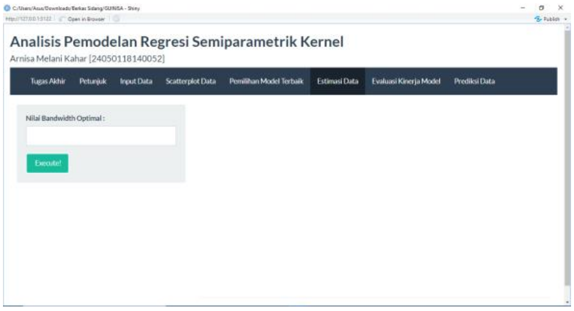
Gambar 2. Tampilan Awal GUI

Gambar 2 menunjukkan tampilan awal yang akan muncul pertama kali aplikasi GUI dibuka. Berdasarkan kebutuhan pelaksanaan analisis yang dilakukan, GUI yang dibuat terdiri dari tujuh tab panel utama yaitu tab panel ‘Petunjuk’ yang berisi petunjuk penggunaan aplikasi, ‘Input Data’ untuk memasukkan data yang akan digunakan, ‘Scatterplot Data’ untuk menampilkan *scatterplot* antara variabel respon dan prediktor, ‘Pemilihan Model Terbaik’ untuk hasil analisis dalam menentukan model terbaik, ‘Estimasi Data’ untuk menampilkan hasil estimasi menggunakan model terbaik, ‘Evaluasi Kinerja Model’ untuk menampilkan hasil perhitungan kebaikan model dan evaluasi kinerja model, dan ‘Prediksi Data’ untuk menampilkan hasil prediksi menggunakan model terbaik. Tampilan dari ketujuh tab panel dapat dilihat pada Gambar 3.



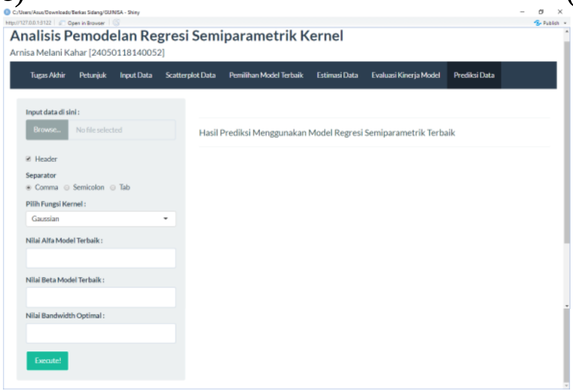
(a) (b)

(c) (d)



(e)

(f)



(g)

Gambar 3. Tampilan Tab Panel (a) Petunjuk, (b) Input Data, (c) Scatterplot Data, (d) Pemilihan Model Terbaik, (e) Estimasi Data, (f) Evaluasi Kinerja Model, dan (g) Prediksi Data

5. KESIMPULAN

Dari hasil analisis pemodelan regresi semiparametrik kernel menggunakan fungsi kernel gaussian, epanechnikov, segitiga, dan *uniform* diperoleh model terbaik berdasarkan nilai MSE terkecil yaitu model kernel segitiga dengan nilai MSE sebesar 130908,7797. Persamaan model regresi semiparametrik kernel segitiga adalah sebagai berikut:

$$\hat{y}_i = -2920{,}1440 + 0{,}00611025x_i + \frac{\sum_{j=1}^{84} K\left(\frac{t_i - t_j}{3{,}24}\right) y_j}{\sum_{k=1}^{84} K\left(\frac{t_i - t_k}{3{,}24}\right)}$$

Model regresi semiparametrik dengan fungsi kernel segitiga yang telah diperoleh memiliki nilai R^2 sebesar 97,15% yang berarti model kuat karena berada pada kategori $R^2 > 67\%$ dan nilai MAPE sebesar 6,97% yang berarti hasil peramalan model sangat akurat karena berada pada kategori MAPE < 10%.

Graphical User Interface (GUI) dibuat dengan tampilan yang telah disesuaikan dengan kebutuhan pelaksanaan analisis pemodelan regresi semiparametrik kernel. Terdapat tab panel untuk menampilkan petunjuk penggunaan aplikasi, tab panel untuk memasukkan *file* data yang akan digunakan untuk pelaksanaan analisis, tab panel yang menunjukkan *scatterplot* data, hasil analisis pemilihan model terbaik, hasil evaluasi kinerja model untuk mengetahui kebaikan garis regresi dan tingkat akurasi peramalan, serta tab panel untuk melaksanakan prediksi menggunakan model terbaik yang telah diperoleh.

DAFTAR PUSTAKA

Blanchard, O. 2006. *Macroeconomics 4th Edition.* New Jersey : Pearson Prentice Hall

Carmona, R.A. 2004. *Statistical Analysis of Financial Data in S-PLUS.* New York : Springer-Verlag

Chin, W.W. 1998. *The Partial Least Squares Approach to Structural Equation Modeling.* Modern Methods for Business Research : Hal. 295-336

Fitriani, A., Srinadi, I.G.A.M., dan Susilawati, M. 2015. *Estimasi Model Regresi Semiparametrik Menggunakan Estimator Kernel Uniform (Studi Kasus : Pasien DBD di RS Puri Raharja).* E-Jurnal Matematika Vol. 4 No. 4 : Hal. 176-180

Gujarati, D. 2003. *Basic Econometrics.* New York : Mc-Grawhill

Hardle, W. 1994. *Applied Nonparametric Regression.* Berlin : Humboldt University

Hardle, W., Liang, H., dan Gao, J. 2000. *Partially Linear Models.* Heidelberg : Physica-Verlag

Kurniasih, D., Mariani, S., dan Sugiman. 2013. *Efisiensi Relatif Estimator Fungsi Kernel Gaussian Terhadap Estimator Polinomial Dalam Peramalan USD Terhadap JPY.* Unnes Journal of Mathematics Vol. 2 No. 2 : Hal. 79-84

Lewis, C.D. 1982. *International and Business Forecasting Methods.* London : Butterworths

Makridakis, S.G., Wheelwright, S.C., dan Hyndman, R.J. 1997. *Forecasting : Methods and Applications.* New York : John Wiley & Sons

Nanda, D.A., Suparti, dan Hoyyi, A. 2016. *Analisis Pengaruh Jumlah Uang Beredar dan Nilai Tukar Rupiah Terhadap Indeks Harga Saham Gabungan Menggunakan Pemodelan Regresi Semiparametrik Kernel.* Jurnal Gaussian Vol. 5 No. 3 : Hal. 373-382

Perdina, N.P. 2012. *Pendugaan Model Regresi Semiparametrik Menggunakan Penduga Kernel.* Bali : Universitas Udayana

Sukirno, S. 2002. *Pengantar Teori Makroekonomi Edisi Kedua.* Jakarta : PT Raja Grafindo Persada

Sunariyah. 2003. *Pengantar Pengetahuan Pasar Modal*. Jakarta : Erlangga

Suparti, Santoso, R., Prahutama, A., dan Devi, A.R. 2018. *Regresi Nonparametrik.* Ponorogo : Wade Group

Wand, M.P. dan Jones, M.C. 1995. *Kernel Smoothing.* New York : Chapman & Hall