

PEMODELAN HARGA EMAS DUNIA MENGGUNAKAN METODE NONPARAMETRIK POLINOMIAL LOKAL DILENGKAPI GUI R

Jody Hendrian¹, Suparti², Alan Prahutama³

^{1,2,3}Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro
jodyhendrian1@gmail.com

ABSTRACT

Investing in gold is a flexible choice because it can be sold at any time and used as an emergency fund. Investors should have the knowledge to predict data from time to time to achieve investment goals. One of the statistical methods for time series data modeling is ARIMA. The ARIMA model is strict with the assumptions that the data must be stationary, the residuals must be normally distributed, independent, and with constant variance, so an alternative model is proposed, namely nonparametric regression model, which has no modeling assumptions requirement. In this study, the daily world gold price data will be modeled using a local polynomial nonparametric model as an alternative because the assumptions in the ARIMA are not fulfilled. The data is divided into 2 parts, namely in sample data from January 2, 2020 to November 30, 2020 to form a model and out sample data from December 1, 2020 to December 31, 2020 used for evaluation of model performance based on MAPE values. The chosen best model is the local polynomial model with Gaussian kernel function of degree 5, bandwidth of 373, and local point of 1744 with an MSE value of 482.6420. The local polynomial model out sample data MAPE value is 0.61%, indicating that the model has excellent forecasting capability. In this study, Graphical User Interface (GUI) using R software with the help of shiny package is also built, making data analyzing easier and generating more interactive display output.

Keywords: Gold, Local Polynomial, ARIMA, MSE, GUI

1. PENDAHULUAN

Investasi emas menjadi pilihan yang fleksibel karena emas dapat dijual kapan saja dan menjadi tabungan darurat. Harga emas cenderung stabil dan naik, terkadang harga emas juga fluktuatif, sifat fluktuatif ini yang harus diperhatikan oleh investor emas. Pengetahuan yang harus dimiliki investor dalam berinvestasi salah satunya dapat membaca pergerakan data dari waktu ke waktu agar tujuan investasi dapat tercapai. Dikutip dari CNN Indonesia (2020), Lukman Leong yang merupakan analis Asia Valbury Futures menilai harga emas terus naik karena pelaku pasar semakin khawatir dengan prospek ekonomi. Pasalnya, jumlah kasus *Covid-19* terus bertambah.

Salah satu analisis stastistika yang dapat digunakan untuk pemodelan adalah model ARIMA. Namun pada pemodelan ARIMA terdapat asumsi-asumsi yang harus dipenuhi yaitu normalitas residual, nonautokorelasi residual, dan homoskedastisitas residual. Sehingga digunakan model alternatif regresi nonparametrik yang tidak terikat asumsi pemodelan.

Salah satu metode regresi nonparametrik yaitu polinomial lokal. Polinomial lokal mempunyai beberapa kelebihan diantaranya adalah kemampuannya dalam beradaptasi

terhadap data yang berarti membagi data ke dalam wilayah tertentu kemudian melakukan estimasi terhadap wilayah yang telah ditetapkan (Fan & Gijbels, 1996). Metode optimasi untuk mencari model polinomial lokal terbaik dapat menggunakan *Mean Square Error* (MSE).

Perhitungan regresi nonparametrik polinomial lokal termasuk rumit dan banyak prosesnya, dibutuhkan *software* untuk menyelesaikannya, salah satunya adalah *software* R. Salah satu pengembangan dalam penelitian ini adalah disusun *Graphical User Interface* (GUI) program R untuk memudahkan analisis data dan menampilkan output yang lebih menarik.

2. TINJAUAN PUSTAKA

2.1. Harga Emas Dunia

Harga emas *London Bullion Market Association* (LBMA) menjadi standar harga emas yang digunakan seluruh dunia dengan sistem *London Gold Fixing*. LBMA menentukan harga emas dua kali dalam sehari setiap hari kerja, yaitu pukul 10.30 (*Gold A.M*) dan pukul 15.00 (*Gold P.M*) dengan mata uang Dolar Amerika Serikat per *troy ounce* yang sama dengan 31,103 gram oleh lima anggota LBMA yaitu Bank of Nova Scotia, Barclays Capital, Deutsche Bank, HSBC, dan Societe Generale. Harga yang digunakan sebagai patokan harga emas seluruh negara di dunia adalah harga *Gold P.M*. (Izati et al., 2019).

2.2. Analisis Regresi

Analisis regresi merupakan sebuah metode statistika yang digunakan untuk melihat hubungan antara variabel respon dengan satu atau lebih variabel prediktor, dimana bentuk kurva atau fungsi regresinya diketahui. Model regresi linier dapat dituliskan dalam bentuk persamaan sebagai berikut: (Montgomery & Runger, 2003)

$$y_i = \sum_{j=0}^k \beta_j x_{ij} + \varepsilon_i \quad i = 1, 2, \dots, n ; \quad j = 0, 1, 2, \dots, k \quad (1)$$

dengan

y_i : variabel respon pengamatan ke- i

x_{ij} : variabel prediktor ke- j pada pengamatan ke- i

β_j : parameter regresi variabel prediktor ke- j

β_0 : intersep (konstanta)

ε_i : residual pengamatan ke- i

n : banyak pengamatan

k : banyak variabel prediktor dalam model regresi

2.3. Regresi Nonparametrik Polinomial Lokal

Regresi polinomial lokal adalah suatu metode regresi nonparametrik, fungsi regresi $f(x)$ ditaksir menggunakan bentuk fungsi polinomial. Fungsi polinomial lokal mengadopsi dari ekspansi deret Taylor disekitaran x_0 . Jika $f(x)$ berderajat m , maka ekspansi dari deret Taylor untuk x yang berada di sekitaran x_0 sebagai berikut:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(m)}(x_0)}{m!}(x - x_0)^m \quad (2)$$

Model regresi polinomial lokal dengan m adalah derajat polinomial dimodelkan sebagai berikut: (Fan & Gijbels, 1996)

$$y_i = f(x_i) + \varepsilon_i \quad , \quad i = 1, 2, \dots, n \quad (3)$$

dengan $f(x_i) = \sum_{r=0}^m \beta_r (x_i - x_0)^r$, $\beta_r = \frac{f^{(r)}(x_0)}{r!}$ untuk $r = 0, 1, \dots, m$

Jika model (3) ditulis dalam bentuk matriks, maka dapat dituliskan sebagai berikut:

$$Y = X\beta + \varepsilon \quad (4)$$

Dengan $X = \begin{pmatrix} 1 & (x_1 - x_0) & (x_1 - x_0)^2 & \dots & (x_1 - x_0)^m \\ 1 & (x_2 - x_0) & (x_2 - x_0)^2 & \dots & (x_2 - x_0)^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & (x_n - x_0) & (x_n - x_0)^2 & \dots & (x_n - x_0)^m \end{pmatrix}$ $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$; $\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}$ dan $\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$

2.4. Estimasi Model Regresi Polinomial Lokal

Estimasi model regresi polinomial lokal dapat menggunakan WLS (*Weighted Least Square*) sehingga diperlukan pembobot. Pembobotan yang dapat digunakan untuk mendapatkan estimasi polinomial lokal adalah fungsi kernel (Eubank, 1988). Fungsi kernel K dengan $bandwidth h$ didefinisikan sebagai berikut:

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right); \quad -\infty < x < \infty \text{ dan } h > 0 \quad (5)$$

Terdapat beberapa jenis fungsi kernel: (Hardle, 1991)

1. Kernel Uniform

$$K(x) = \begin{cases} \frac{1}{2}, & \text{untuk } |x| \leq 1 \\ 0, & \text{untuk } x \text{ yang lain} \end{cases}$$

2. Kernel Segitiga

$$K(x) = \begin{cases} 1 - |x|, & \text{untuk } |x| \leq 1 \\ 0, & \text{untuk } x \text{ yang lain} \end{cases}$$

3. Kernel Epanechnikov

$$K(x) = \begin{cases} \frac{3}{4}(1 - x^2), & \text{untuk } |x| \leq 1 \\ 0, & \text{untuk } x \text{ yang lain} \end{cases}$$

4. Kernel Gaussian

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \text{ untuk } |x| < \infty$$

Menurut Hardle dan Linton (1994), fungsi kernel merupakan fungsi yang riil, kontinu, terbatas, dan simetris, serta memenuhi $\int_{-\infty}^{\infty} K(x) dx = 1$

Parameter β pada persamaan (3) bergantung pada titik x_0 yang disebut sebagai titik lokal. Parameter β diestimasi menggunakan *Weighted Least Square* dengan fungsi kernel sebagai pembobotnya. Berikut fungsi yang diminimumkan:

$$RSS = \sum_{i=1}^n \{y_i - \sum_{j=0}^m \beta_j (x_i - x_0)^j\}^2 K\left(\frac{x_i - x_0}{h}\right) \quad (6)$$

dengan h adalah *bandwidth* yang mengontrol ukuran persekitaran titik lokal x_0 . Persamaan (6) dapat dituliskan dalam bentuk matriks menjadi:

$$RSS = (Y - X\beta)^T W (Y - X\beta) \quad (7)$$

dengan $W = \text{diag} \left\{ K\left(\frac{x_i - x_0}{h}\right) \right\}$ diperoleh estimasi model:

$$\hat{Y} = X\hat{\beta} = X(X^T W X)^{-1} X^T W Y \quad (8)$$

Hal yang harus diperhatikan dalam memilih model polinomial lokal adalah pemilihan *bandwidth* h . Pemilihan *bandwidth* yang terlalu besar mengakibatkan plot hasil estimasi model akan menjauhi plot data awal sehingga menjadi sangat halus (*oversmoothing*), bias pemodelannya akan besar, dan keragaman akan kecil. Pemilihan *bandwidth* yang terlalu kecil mengakibatkan plot hasil estimasi model yang berliuk-liuk (*undersmoothing*), bias pemodelannya kecil, dan keragaman akan besar. Selain *bandwidth*, pemilihan derajat polinomial lokal juga harus diperhatikan. Derajat yang besar akan mengurangi bias pemodelan tapi menyebabkan keragaman yang besar. Oleh karena itu, harus dicari *bandwidth* dan derajat polinomial yang optimal sehingga dapat menyeimbangkan bias dan keragaman agar diperoleh estimasi yang baik (Suparti & Prahutama, 2016).

2.5. Pemilihan Model Polinomial Lokal Terbaik

Model terbaik polinomial lokal merupakan model dengan *bandwidth*, titik lokal, bobot fungsi kernel, dan derajat polinomial yang optimal. Metode optimasi yang digunakan pada penelitian ini adalah meminimalkan nilai *Mean Square Error* (MSE). MSE adalah rata-rata dari kuadrat residual model yang dirumuskan sebagai berikut:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (9)$$

Koefisien determinasi dapat digunakan untuk memilih model terbaik yang menunjukkan seberapa besar persentase keragaman dalam variabel respon untuk dijelaskan oleh variabel prediktor. Koefisien determinasi dapat dirumuskan sebagai berikut :

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (10)$$

Nilai R^2 berada diantara 0 sampai 1, jika nilai R^2 sama dengan 1 maka variabel respon dapat dijelaskan oleh variabel prediktor sepenuhnya, jika nilai R^2 sama dengan 0 maka

variabel respon tidak dapat dijelaskan oleh variabel prediktor. Kriteria nilai R^2 dapat dibagi menjadi tiga tingkatan yaitu 0,75; 0,50; dan 0,25 sebagai model kuat, moderat, dan lemah (Hair et al., 2017).

2.6. Model ARIMA

Model ARIMA (*Autoregressive Integrated Moving Average*) adalah model parametrik yang terdiri dari AR, MA, ARMA, dan ARIMA.

1. Model Autoregresif (AR)

Bentuk umum proses Autoregresif orde p: (Soejoeti, 1987)

$$Z_i = \phi_1 Z_{i-1} + \phi_2 Z_{i-2} + \dots + \phi_p Z_{i-p} + a_i \quad (10)$$

2. Model Moving Average (MA)

Bentuk umum proses Moving Average orde q: (Wei, 1989)

$$Z_i = a_i - \theta_1 a_{i-1} - \theta_2 a_{i-2} - \dots - \theta_q a_{i-q} \quad (11)$$

3. Model ARMA(p,q) (Autoregresif orde p dan Moving Average orde q)

Bentuk umum model campuran AR orde p dan MA orde q:

$$Z_i = \phi_1 Z_{i-1} + \phi_2 Z_{i-2} + \dots + \phi_p Z_{i-p} + a_i - \theta_1 a_{i-1} - \theta_2 a_{i-2} - \dots - \theta_q a_{i-q} \quad (12)$$

4. Model ARIMA(p,d,q)

Bentuk umum model yang melibatkan proses AR(p), MA(q), dan tingkat *differencing* (d):

$$\phi_p(B)(1 - B)^d Z_i = \theta_q(B) a_i \quad (13)$$

dengan B adalah operator *backward shift*, $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$, $\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$

5. Model Subset ARIMA

Model ARIMA yang tergeneralisasi dan tidak dapat dinyatakan dalam bentuk umum model disebut model subset ARIMA (Tarno, 2013). Contoh model subset ARIMA adalah ARIMA([2,5],0,[12]) dapat dirumuskan sebagai berikut:

$$(1 - \phi_2 B^2 - \phi_5 B^5) Z_i = (1 - \theta_{12} B^{12}) a_i \quad (14)$$

Identifikasi model ARIMA dapat menggunakan plot *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF). Pemilihan model terbaik ARIMA yang teridentifikasi diawali dengan uji asumsi yang terdiri dari uji normalitas residual, nonautokorelasi residual, dan homoskedastisitas residual. Jika terdapat asumsi yang tidak terpenuhi, maka model ARIMA tidak dapat digunakan.

2.7. Model Runtun Waktu Nonparametrik

Jika pemodelan ARIMA tidak dapat digunakan, maka model nonparametrik dapat digunakan sebagai alternatif pemodelan. Dalam penggunaan model runtun waktu nonparametrik data dimodelkan berdasarkan signifikan lag pada *Partial Autocorrelation Function* (PACF). Jika data signifikan pada lag ke-p maka model runtun waktu adalah

AR(p). Mengacu pada persamaan (3) dan (10) maka model regresi polinomial lokal menjadi:

$$Z_i = f(Z_{i-p}) + \varepsilon_i, \quad i = (p + 1), (p + 2), \dots, n \quad (15)$$

dengan $f(Z_{i-p}) = \sum_{r=0}^m \beta_r (Z_{i-p} - x_0)^r$, $\beta_r = \frac{f^{(r)}(x_0)}{r!}$ untuk $r = 0, 1, \dots, m$

2.8. Evaluasi Kinerja Model

Evaluasi kinerja model menggunakan nilai MAPE (*Mean Absolute Percentage Error*) data *out sample*. MAPE yaitu persentase rata-rata nilai absolut *error* dibagi data aktualnya yang dirumuskan sebagai berikut:

$$MAPE = \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \quad (25)$$

Nilai MAPE adalah $MAPE < 10\%$ berarti model memiliki kemampuan peramalan yang sangat baik, $10\% \leq MAPE < 20\%$ berarti model memiliki kemampuan peramalan yang baik, $20\% \leq MAPE < 50\%$ berarti model memiliki kemampuan peramalan cukup baik, dan $MAPE > 50\%$ berarti model memiliki kemampuan peramalan yang buruk (Chang et al., 2007).

2.9. Graphical User Interface (GUI) Program R

Pembuatan *Graphical User Interface* pada penelitian ini menggunakan paket *Shiny* program R. *Shiny* terdiri dari tiga komponen yaitu *User Interface* (UI), *Server*, dan *ShinyApp*. Pada bagian *User Interface* bermanfaat untuk panel kontrol, pemasukan nilai input, dan penyajian output. Panel kontrol yaitu panel untuk mengontrol input berupa data, variabel, model, tergantung kompleksitas model. *Server* pada *shiny* merupakan otak dari program yang bertugas melakukan simulasi, berbagai analisis data sesuai pilihan pengguna dan selanjutnya mengirim hasilnya ke bagian output (Tirta, 2014).

3. METODE PENELITIAN

3.1. Sumber Data dan Variabel Penelitian

Data yang digunakan adalah data historis harga emas dunia bersumber dari *website investing.com* yang standar perhitungannya menggunakan *London Gold Fixing* dalam mata uang Dolar Amerika per *troy ounce*. Data dibagi menjadi dua, yaitu data *in sample* dari tanggal 2 Januari 2020 sampai 30 November 2020 dan data *out sample* dari 1 Desember 2020 sampai 31 Desember 2020. Variabel yang digunakan dalam penelitian ini adalah harga emas harian periode 2 Januari sampai 31 Desember 2020 pada waktu ke- i atau (Z_i) sebagai variabel respon, sedangkan variabel prediktor adalah data pada waktu ke- $(i - p)$ atau Z_{i-p} berdasarkan lag signifikan.

3.2. Langkah-langkah Analisis

Dalam penelitian ini menggunakan software R 3.6.2 dengan bantuan GUI R dari paket *Shiny*. Adapun tahapan analisis data sebagai berikut:

1. Menyiapkan aplikasi GUI R sesuai dengan kebutuhan analisis.
2. Menentukan data *in sample* dan *out sample* dan melakukan eksplorasi data.
3. Memodelkan data dengan ARIMA.

4. Cek asumsi pada model ARIMA kemudian menentukan model terbaik dengan *Mean Square Error* (MSE).
5. Menentukan lag maksimum signifikan (p) data *in sample* melalui plot PACF untuk menentukan variabel prediktor pada model nonparametrik polinomial lokal.
6. Membuat plot data X_{i-p} dan X_i yang ditentukan berdasarkan lag signifikan pada PACF.
7. Menghitung nilai MSE pada kombinasi model derajat 1 sampai 5 dengan pembobot fungsi kernel Gaussian, Uniform, dan Segitiga, serta nilai *bandwidth* dan titik lokal yang ditentukan.
8. Memilih model terbaik polinomial lokal dengan nilai MSE minimum.
9. Membandingkan model ARIMA dengan polinomial lokal dan memilih model terbaik berdasarkan nilai MSE terkecil.
10. Menentukan kinerja model terbaik dengan MAPE *out sample*.

4. HASIL DAN PEMBAHASAN

4.1. Pembuatan dan Penggunaan GUI R

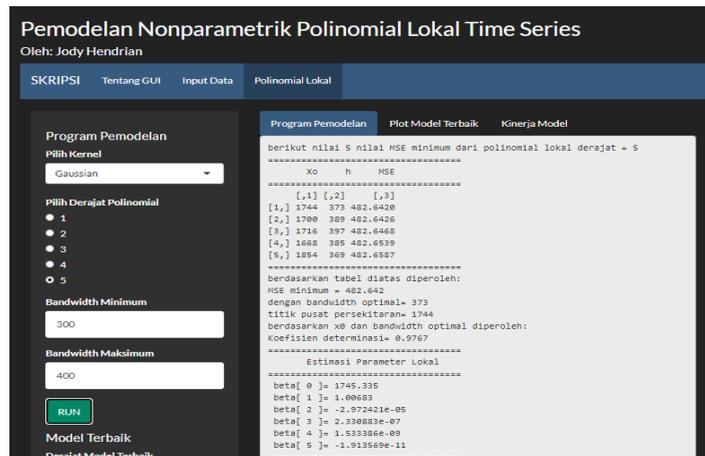
Pembuatan GUI R meliputi bagian *user interface* yang mengatur penempatan input dan output menggunakan perintah `ui<-fluidPage()`. Tiap input yang dilakukan menggunakan bantuan perintah-perintah pada UI harus diberi identitas. Proses pengolahan yang dijalankan oleh GUI akan didefinisikan dalam fungsi *shinyApp* yaitu pada *server*. Berikut tampilan program UI dan *server*:

```

+ }},
+ tabPanel("Input Data",
+   sidebarLayout(
+     sidebarPanel(
+       fileInput("datain","Masukkan data timeseries in sample",accept = c("text","txt")),
+       fileInput("data","Masukkan data in sample X(i) dan X(i-p) berdasarkan lag signifikan PACF",accept = c("text","txt")),
+       fileInput("dataout","Masukkan data out sample",accept = c("text","txt")),
+     ),
+     mainPanel(
+       tabsetPanel(type = "pills",id = "navbar",
+         tabPanel("Plot",
+           plotOutput("pacfplot"),
+           plotOutput("pacfplot2"),
+           "Value=Plot"),
+         tabPanel("Data",verbatimTextOutput("deskriptif"),tableOutput("inputdata"),tableOutput("inputdata2")),
+       )
+     )
+   ),
+   tabPanel("Polinomial Lokal",
+     sidebarLayout(
+       sidebarPanel(
+         selectInput("kernel","Pilih Kernel",choices = c("Gaussian","Uniform","Segitiga")),
+         radioButtons("p","Pilih Derajat Polinomial",list("1","2","3","4","5")),
+         textInput("band","Bandwidth Minimum"),
+         textInput("width","Bandwidth Maksimum"),
+         actionButton("hitung1","RUM",class="btn-success"),
+         textInput("p","Derajat Model Terbaik"),
+         textInput("h","Bandwidth Model Terbaik"),
+         textInput("k","pnp Model Terbaik"),
+         actionButton("hitung2","RUM",class="btn-success")
+       ),
+       mainPanel(
+         tabsetPanel(type = "pills",id = "navbar",
+           tabPanel("Program Pemodelan",verbatimTextOutput("mse"),value="Program Pemodelan"),
+           tabPanel("Plot Model Terbaik",plotOutput("polfit"),plotOutput("plotin"),value="Plot Model Terbaik"),
+         )
+       )
+     )
+   ),
+   Y<-file[,2]
+   pc<-as.numeric(inputP)
+   band<-as.numeric(inputBand)
+   width<-as.numeric(inputWidth)
+   if(input$kernel=="Gaussian")
+   {
+     output$mae<-renderPrint((mse(X,Y,band,width,p,Gaussian)))
+   }
+   if(input$kernel=="Segitiga")
+   {
+     output$mae<-renderPrint((mse(X,Y,band,width,p,Segitiga)))
+   }
+   if(input$kernel=="Uniform")
+   {
+     output$mae<-renderPrint((mse(X,Y,band,width,p,Uniform)))
+   }
+   })
+   observeEvent(input$hitung2,{
+     data3<-input$data
+     if(is.null(input$data)) {return()}
+     file<-read.csv(input$data$datapath,header = T ,sep ='\t')
+     X<-file[,1]
+     Y<-file[,2]
+     data2<-input$dataout
+     if(is.null(input$dataout)) {return()}
+     file1<-read.csv(input$dataout$datapath,header = T ,sep ='\t')
+     X1<-file1[,1]
+     Y1<-file1[,2]
+     pc<-as.numeric(inputP)
+     k0<-as.numeric(inputK0)
+     h<-as.numeric(inputH)
+     if(input$kernel=="Gaussian")
+     {
+       output$polfit<-renderPlot((polynomialFit(X,Y,p,k0,h,Gaussian)))
+       output$plotin<-renderPlot((plotin(X,Y,p,k0,h,Gaussian)))
+       output$outsample<-renderPrint((kinerja(X,Y,X1,Y1,p,k0,h,Gaussian)))
+       output$plotout<-renderPlot((kinerja(X,Y,X1,Y1,p,k0,h,Gaussian)))
+     }
  
```

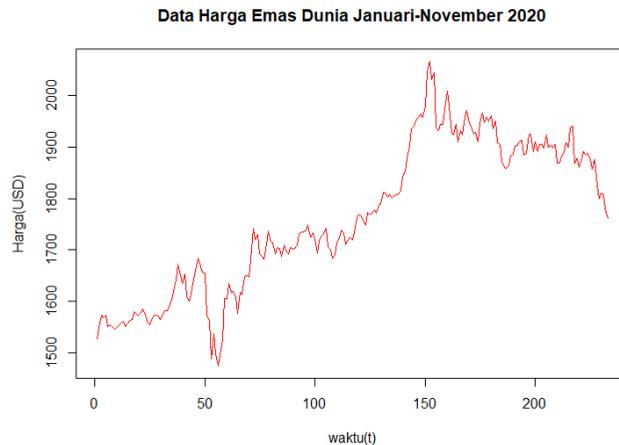
Gambar 1 Tampilan Program UI dan Server

Setelah objek UI dan *server* selesai dibentuk maka langkah selanjutnya yaitu *running* aplikasi dengan fungsi `shinyApp(ui,server)` dan aplikasi dapat dipakai. Berikut tampilan GUI untuk pencarian model polinomial lokal:



Gambar 2 Tampilan GUI R Polinomial Lokal

4.2. Eksplorasi Data



Gambar 3 Plot Data Runtun Waktu Harga Emas *In Sample*

Berdasarkan Gambar 3, terlihat data mengalami fluktuasi pada pertengahan bulan Februari 2020 sampai dengan 1 April 2020 yaitu data ke 34 sampai 65. Kemudian naik sampai puncaknya pada 6 Agustus 2020 yaitu data ke 152 mencapai USD 2067,15 per *trou ounce* dan cenderung turun sampai bulan November 2020. Statistik deskriptif data dapat dilihat pada Tabel 6 berikut ini:

Tabel 1 Statistik Deskriptif Data *In Sample*

Rata-rata	Maksimum	Minimum	Varian
1762,547	2067,15	1474,25	20910,99

Berdasarkan Tabel 6, rata-rata data harga emas harian periode Januari 2020 sampai November 2020 sebesar 1762,547 dengan nilai maksimum 2067,15, minimum 1474,25, dan varian data sebesar 20910,99.

4.3. Pemodelan ARIMA

Sebelum melakukan pemodelan data menggunakan ARIMA, data dicek terlebih dahulu kestasionerannya dalam mean dan varian. Nilai lambda diperoleh 1 maka data stasioner dalam varian dan diuji stasioner dalam mean menggunakan *Augmented Dickey Fuller* diperoleh kesimpulan data harus di-*differencing* satu kali karena data asli tidak stasioner dalam mean. Model ARIMA yang diperoleh adalah model subset ARIMA([4],1,[4]), ARIMA([21],1,[4]), ARIMA([4,21],1,[4]), ARIMA([4],1,0), ARIMA([21],1,0), ARIMA([4,21],1,0), dan ARIMA(0,1,[4]). Seluruh model tidak memnuhi asumsi normalitas residual dan homoskedastisitas residual. Oleh karena itu, model ARIMA tidak dapat digunakan untuk pemodelan harga emas pada penelitian ini. Sebagai alternatif digunakan model nonparametrik runtun waktu polinomial lokal.

4.4. Pemodelan Regresi Nonparametrik Polinomial Lokal

Berdasarkan plot PACF data in sample signifikan pada lag 1 maka pemodelan regresi nonparametrik polinomial lokal mengambil variabel prediktor Z_{i-1} dan respon Z_i . Pencarian model terbaik dilakukan dengan *trial-error* dari sekumpulan nilai titik lokal (x_0) yang terdapat dalam interval tertentu yang dicobakan. Diketahui data minimum 1474,25 dan data maksimum 2067,15 sehingga nilai titik lokal dapat dicobakan pada interval 1475 sampai dengan 2067, *bandwidth* dicobakan mulai dari 20 sampai 400 dan derajat polinomial dicoba derajat 1,2,3,4, dan 5. Sedangkan untuk pembobot kernelnya dicobakan kernel Gaussian, Uniform, dan Segitiga. Pemilihan model polinomial lokal terbaik digunakan optimasi *Mean Square Error* (MSE).

Terpilih model terbaik adalah model dengan fungsi pembobot kernel Gaussian, nilai *bandwidth* 373 dan titik lokal 1744. Model tersebut memiliki nilai R-Square yang tinggi yaitu 0,9767 atau 97,67% termasuk dalam kriteria model yang sangat kuat. Berikut hasil estimasi modelnya:

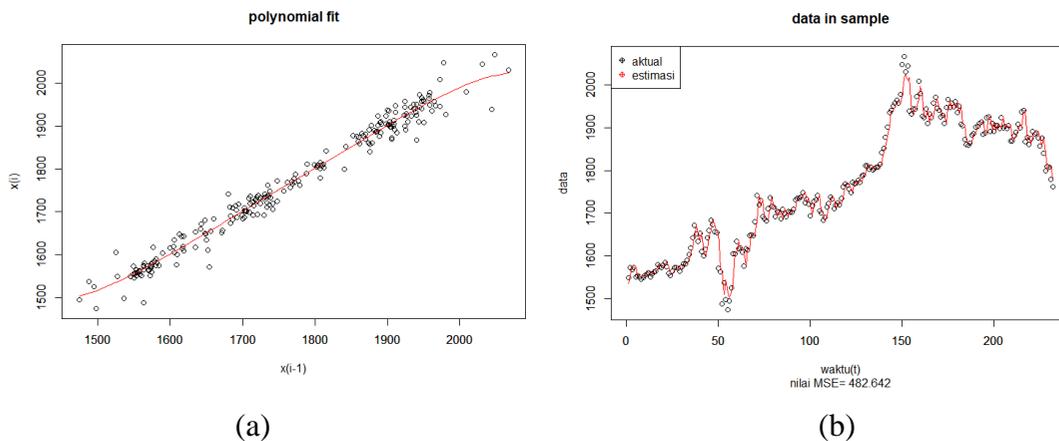
Tabel 2 Estimasi Model Polinomial Lokal Derajat 5

Variabel	Hasil Estimasi
$\hat{\beta}_0$	1745,335
$\hat{\beta}_1$	1,00683
$\hat{\beta}_2$	-0,00002972421
$\hat{\beta}_3$	$2,330883 \times 10^{-7}$
$\hat{\beta}_4$	$1,533386 \times 10^{-9}$
$\hat{\beta}_5$	$-1,913569 \times 10^{-11}$
x_0	1744

Model polinomial lokal runtun waktu yang terbentuk adalah:

$$\hat{Z}_i = 1745,335 + 1,00683(Z_{i-1} - 1744) - 0,00002972421(Z_{i-1} - 1744)^2 + 2,330883 \times 10^{-7}(Z_{i-1} - 1744)^3 + 1,533386 \times 10^{-9}(Z_{i-1} - 1744)^4 - 1,913569 \times 10^{-11}(Z_{i-1} - 1744)^5$$

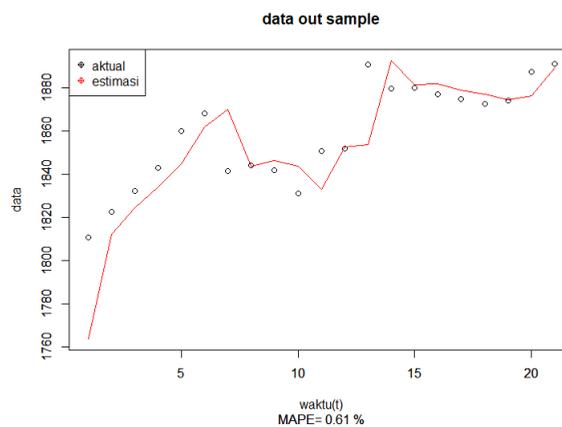
Plot data modifikasi yaitu Z_i dan Z_{i-1} terhadap nilai estimasinya dan plot data asli terhadap estimasinya sebagai berikut:



Gambar 4 (a) Plot Data Modifikasi dan Estimasinya
(b) Plot Data Asli dan Estimasinya

Berdasarkan Gambar 4, secara visual plot data modifikasi dan data asli dapat didekati oleh data estimasinya dengan baik. Hal tersebut membuktikan bahwa model polinomial lokal dapat digunakan dalam pemodelan harga emas ini.

4.5. Evaluasi Kinerja Model Terbaik



Gambar 5 Plot Data Aktual *Out Sample* dan Estimasinya

Model polinomial lokal menghasilkan plot estimasi yang baik untuk data *in sample*. Berdasarkan Gambar 5, secara visual plot data aktual *out sample* dapat didekati dengan baik oleh data estimasinya dan nilai MAPE *out sample* sebesar 0,61%. Nilai MAPE kurang dari 10% menunjukkan model memiliki kemampuan peramalan yang sangat baik. Hal tersebut membuktikan bahwa model polinomial lokal derajat 5 memiliki kemampuan sangat baik secara *in sample* dan *out sample*.

5. KESIMPULAN

Berdasarkan analisis dan pembahasan yang telah dibahas sebelumnya dapat disimpulkan sebagai berikut:

1. Model ARIMA yang merupakan model parametrik terdapat asumsi yang tidak terpenuhi yaitu normalitas residual dan homoskedastisitas residual, maka model nonparametrik runtun waktu polinomial lokal menjadi alternatif pemodelan.
2. Model nonparametrik polinomial lokal terbaik adalah model derajat 5 fungsi pembobot kernel Gaussian dengan *bandwidth* $h = 373$, dan titik lokal $x_0 = 1744$.
3. Pada evaluasi kinerja model polinomial lokal menggunakan MAPE *out sample* yang dihitung berdasarkan data aktual *out sample* dan data estimasinya menghasilkan nilai MAPE sebesar 0,61% yang mana termasuk dalam kategori model yang memiliki kemampuan peramalan yang sangat baik.
4. Pembuatan GUI R menjadi salah satu pengembangan penelitian sebelumnya yang hanya memakai *software* R berbasis *Command Line Interface* (CLI). Analisis data menggunakan GUI R dapat menghasilkan output yang lebih menarik dan lebih memudahkan pengguna dalam menganalisis data.

DAFTAR PUSTAKA

- Chang, P.-C., Wang, Y.-W., & Liu, C.-H. 2007. *The Development of a Weighted Evolving Fuzzy Neural Network for PCB Sales Forecasting*. Expert Systems with Application, 32(88–89).
- Eubank, R. L. 1988. *Nonparametric Regression and Spline Smoothing*. New York: Marcell dekker, Inc.
- Fan, J., & Gijbels, I. 1996. *Local Polynomial Modelling and Its applications: Monographs on Statistics and Applied Probability 66*. London: Chapman and Hall.
- Hair, J. F., Hult, G., Ringle, C. M., & Sartetd, M. 2017. *A Primer on Partial Least Squares Structural Equation Model (PLS-SEM)* (2nd ed.). USA: SAGE Publications, Inc.
- Hardle, W. 1991. *Smoothing Techniques with Implementation in S*. New York: Springer Verlag.
- Hardle, W. & Linton, O. 1994. *Applied Nonparametric Methods*. Handbook of Econometrics, 9(2), 2295-2339
- Izati, N. A., Warsito, B., & Widiharih, T. 2019. Prediksi Harga Emas Menggunakan Feed Forward Neural Network dengan Metode Extreme Learning Machine. *Jurnal Gaussian*, 8(2), 171–183.
- Montgomery, D. C., & Runger, G. C. 2003. *Applied Statistics and Probability for Engineers*. New Jersey: John Wiley & Sons, Inc.
- Soejoeti, Z. 1987. *Analisis Runtun Waktu*. Karunia Jakarta.
- Suparti, & Prahutama, A. 2016. *Pemodelan Regresi Nonparametrik Menggunakan Pendekatan Polinomial Lokal Pada Beban Listrik Di Kota Semarang*. Media Statistika, 9(2), 85–93.
- Tarno. 2013. *Kombinasi Prosedur Pemodelan Subset ARIMA dan Deteksi Outlier untuk Data Runtun Waktu*. Prosiding Seminar Nasional Statistika Universitas Diponegoro. 583–592.

Tirta, I. M. 2014. *Pengembangan E-Modul Statistika Terintegrasi dan Dinamik dengan R-Shiny dan mathJax*. Prosiding Seminar Nasional Matematika, 223–232.

Wei, W. W. .1989. *Time Series Analysis: Univariate and Multivariate Methods*. Canada: Addison Wesley Publishing Company.