

**PENDUGAAN REGRESI NONPARAMETRIK  
DENGAN FUNGSI KERNEL GAUSSIAN**

**SKRIPSI**

oleh:  
**ERWIN ANDRIANTO**  
**0910950035-95**



**PROGRAM STUDI STATISTIKA  
JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS BRAWIJAYA  
MALANG  
2014**

**LEMBAR PENGESAHAN SKRIPSI**

**PENDUGAAN REGRESI NONPARAMETRIK  
DENGAN FUNGI KERNEL GAUSSIAN**

oleh:

**ERWIN ADNRIANTO**  
**0910950035-95**

**Setelah dipertahankan di depan Majelis Penguji  
pada tanggal 08 Agustus 2014  
dan dinyatakan memenuhi syarat untuk memperoleh gelar  
Sarjana Sains dalam bidang Statistika**

**Dosen Pembimbing**

**Dr. Ir. Atiek Iriany, MS**  
**NIP. 196308091988022001**

**Mengetahui,**  
**Ketua Jurusan Matematika**  
**Fakultas MIPA Universitas Brawijaya**

**Dr. Abdul Rouf Alghofari, M.Sc**  
**NIP. 196709071992031001**

## LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Erwin Andrianto  
NIM : 0910950035-95  
Jurusan : Matematika  
Program Studi : Statistika  
Skripsi berjudul : PEDUGAAN REGRESI  
NONPARAMETRIK DENGAN  
FUNGSI KERNEL GAUSSIAN

Dengan ini menyatakan bahwa:

1. Isi dari Skripsi yang saya buat adalah benar-benar karya sendiri dan tidak menjiplak karya orang lain, selain nama-nama yang termaktub di isi dan tertulis di daftar pustaka dalam Skripsi ini.
2. Apabila dikemudian hari ternyata Skripsi yang saya tulis terbukti hasil jiplakan, maka saya akan bersedia menanggung segala resiko yang akan saya terima.

Demikian pernyataan ini dibuat dengan segala kesadaran.

Malang, 08 Agustus 2014  
Yang menyatakan,

Erwin Andrianto  
NIM. 0910950035-95

# PENDUGAAN REGRESI NONPARAMETRIK DENGAN FUNGI KERNEL GAUSSIAN

## ABSTRAK

Analisis regresi nonparametrik merupakan metode pendugaan kurva regresi yang digunakan jika tidak ada informasi sebelumnya tentang bentuk kurva regresi atau tidak terikat pada asumsi bentuk fungsi tertentu. Estimasi fungsi regresi nonparametrik dilakukan berdasarkan data pengamatan dengan menggunakan teknik pemulusan (*smoothing*). Penelitian ini bertujuan untuk memperlihatkan pendekatan penduga kernel dalam regresi nonparametrik pada data sekunder berupa kadar monocyte pasien penderita Tubekolosis Paru. Dari hasil perhitungan didapat nilai koefisien determinasi regresi linier sebesar 0.006, hasil perhitungan koefisien determinasi regresi kuadratik sebesar 0.024, hasil perhitungan koefisien determinasi regresi kubik sebesar 0.152, dan hasil perhitungan koefisien determinasi regresi nonparametrik sebesar 0.948. Dari hasil perhitungan dapat disimpulkan bahwa dalam pendugaan kadar monocyte pasien penderita tuberkolosis paru di rumah sakit Dr. Saiful Anwar (RSSA) metode yang lebih tepat adalah dengan menggunakan regresi nonparametrik dengan fungsi kernel.

Kata Kunci: Regresi Nonparametrik, Penduga Kernel,  
*Bandwidth*, GCV

# NONPARAMETRIC REGRESSION ESTIMATION WITH KERNEL GAUSSIAN

## ABSTRACT

Nonparametric regression analysis is a curve estimation methods are used to estimate a regression curve if there is no previous information about the shape of regression curve or not tied to a specific function form assumptions. Estimation of nonparametric regression based on observation data using smoothing technique. This research aims to demonstrate the kernel estimator in nonparametric regression on the secondary data which is monocyte levels of patients with Pulmonary Tuberculosis. Based on calculation we obtained determination coefficient value of linier regression is 0.006, determination coefficient value of quadratic regression is 0.024, determination coefficient value of cubic regression is 0.152, and determination coefficient value of nonparametric regression is 0.948. based on calculation it can be concluded that nonparametric regression with kernel functions is more appropriate than parametric regression on the onocyte levels of patients with pulmonary tuberculosis at hostital Dr. Saiful Anwar (RSSA).

*Key Words: Nonparametric regression, Kernel Estimator, Bandwith, GCV.*

## KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas berkat, rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains dalam bidang Statistika.

Dalam penyusunan skripsi ini, penulis telah banyak dibantu oleh berbagai pihak. Oleh karena itu, dalam kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Dr. Atiek Iriany, M.S selaku dosen pembimbing atas motivasi, waktu dan bimbingan yang telah diberikan.
2. Dr. Ir. Maria Bernadetha M. selaku dosen penguji I atas waktu, ilmu dan bimbingan yang telah diberikan.
3. Samingun Handoyo, S.Si., M.Cs selaku dosen penguji II atas saran dan masukan yang telah diberikan.
4. Dr. Abdul Rouf Alghofari, M.Sc. selaku Ketua Jurusan Matematika Fakultas MIPA Universitas Brawijaya.
5. Ibu, Bapak, dan seluruh keluarga besar tercinta atas segala doa, kasih sayang, dukungan, dan nasihat yang telah diberikan.
6. Teman-teman Prodi Statistika angkatan 2008, 2009 dan 2010 atas kebersamaan dan dukungan. Grup Ginyu dan personil Mertojoyo yang senantiasa menemani susah dan senang dalam pembuatan skripsi ini.
7. Semua pihak yang telah membantu dalam penyusunan skripsi ini.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih jauh dari kesempurnaan mengingat keterbatasan kemampuan penulis. Untuk itu, dengan segala kerendahan hati penulis mengharap kritik dan saran. Semoga skripsi ini dapat bermanfaat bagi semua pihak yang membutuhkan.

Malang, Agustus 2014

Penulis



## DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>HALAMAN PENGESAHAN</b> .....	ii
<b>HALAMAN PERNYATAAN</b> .....	iii
<b>ABSTRAK</b> .....	iv
<b>ABSTRACT</b> .....	v
<b>KATA PENGANTAR</b> .....	vi
<b>DAFTAR ISI</b> .....	vii
<b>DAFTAR GAMBAR</b> .....	ix
<b>DAFTAR TABEL</b> .....	x
<b>DAFTAR LAMPIRAN</b> .....	xi
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	2
1.3 Batasan Masalah .....	2
1.4 Tujuan Penelitian .....	2
1.5 Manfaat Penelitian .....	3
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Analisis Reresi Parametrik .....	5
2.2 Regresi Nonparametrik .....	7
2.2.1 Fungsi Kernel .....	8
2.2.2 Regresi Polinomial Lokal .....	9
2.2.3 Metode Kuadrat Terkecil Terboboti .....	9
2.2.4 Pemilihan <i>Bandwith</i> Optimal .....	11
2.3 Koefisien Determinasi .....	13
2.4 Tuberkulosis .....	14
<b>BAB III METODE PENELITIAN</b>	
3.1 Sumber Data .....	17
3.2 Metode Analisis .....	17
<b>BAB IV HASIL DAN PEMBAHASAN</b>	
4.1 Penduga Kurva Regresi Nonparametrik .....	21
4.2 Penduga Kernel .....	23
4.3 Pemilihan Model Regresi Nonparametrik .....	26
4.4 Pemilihan Regresi Parametrik dengan Regresi Nonparametrik .....	26

**BAB V KESIMPULAN DAN SARAN**

5.1 Kesimpulan.....	29
5.2 Saran.....	29

<b>DAFTAR PUSTAKA.....</b>	<b>31</b>
----------------------------	-----------

<b>LAMPIRAN.....</b>	<b>33</b>
----------------------	-----------





## DAFTAR GAMBAR

	Halaman
Gambar 3.1 Diagram Alir Regresi Nonparametrik .....	19
Gambar 4.1 Plot data antara Kadar Monocyte dengan usia pasien (tahun) .....	23
Gambar 4.2 Model regresi linier kadar monocyte dengan usia pasien (tahun) .....	24
Gambar 4.3 Model regresi kuadratik kadar monocyte dengan usia pasien (tahun) .....	24
Gambar 4.4 Model regresi kubik kadar monocyte dengan usia pasien (tahun) .....	25
Gambar 4.4 Model regresi nonparametrik dengan fungsi kernel gaussian .....	27



## DAFTAR TABEL

	Halaman
Tabel 4.1 Nilai Koefisien Determinasi Model .....	25
Tabel 4.2 Perbandingan Regresi Parametrik dan Regresi Nonparametrik .....	27

UNIVERSITAS BRAWIJAYA



## DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data Pasien Penderita Penyakit Tuberkulosis Paru Rumah Sakit Dr Saiful Anwar (RSSA), Malang.....	33
Lampiran 2. Deskripsi Statistik Data Pasien Penderita Penyakit Tuberkulosis Paru .....	33
Lampiran 3. Makro Program Untuk Menentukan Nilai <i>Bandwidth</i> (h) Optimal dengan Fungsi Kernel Gaussian .....	34
Lampiran 4. Nilai <i>Bandwidth</i> (h) dan GCV dengan Fungsi Kernel Gaussian pada Selang Kenaikan $h=0.05$ .....	38
Lampiran 5. Nilai <i>Bandwidth</i> (h) dan GCV dengan Kernel Gaussian pada selang kenaikan $h=0.005$ .....	40
Lampiran 6. Nilai <i>Bandwidth</i> (h) dan GCV dengan Kernel Gaussian pada selang kenaikan $h=0.001$ .....	41
Lampiran 7. Program Untuk Menentukan Pendugaan $\hat{Y}_i$ dengan Kernel Gaussian .....	42
Lampiran 8. Hasil Pendugaan $\hat{Y}_i$ dengan Kernel Gaussian dengan <i>Bandwidth</i> (h)=0.005.....	45
Lampiran 9. Kadar Monocyte (Y) Dan Hasil Pendugaan Kadar Monocyte ( $\hat{Y}_i$ ).....	47

UNIVERSITAS BRAWIJAYA



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Analisis regresi merupakan metode analisis data yang menghubungkan peubah respon dengan peubah prediktor yang digunakan untuk melihat pengaruh peubah prediktor terhadap peubah respon. Penduga respon ( $Y$ ) dan satu peubah prediktor ( $X$ ) berhubungan secara linier sebagai:

$$Y_i = m(X_i) + \varepsilon_i \quad i = 1, 2, \dots, n$$

di mana  $\varepsilon_i$  adalah sisaan ke- $i$  yang menyebar normal dengan rata-rata 0 dan ragam  $\sigma^2$ . Fungsi  $m(X_i)$  adalah fungsi regresi yang akan diduga, namun tidak diketahui bentuknya.

Pendekatan yang digunakan dalam menduga fungsi regresi, yaitu pendekatan parametrik dan nonparametrik. Dalam pendekatan parametrik, bentuk hubungan antara peubah respon dengan peubah prediktor diketahui (diperkirakan) berdasarkan diagram pencar (linier, kuadratik, eksponensial dan polinomial). Penggunaan regresi parametrik harus memenuhi asumsi yang ketat yaitu sisaan menyebar normal ( $0, \sigma^2$ ). Penyimpangan terhadap asumsi dapat diatasi melalui transformasi sehingga diperoleh model regresi yang sesuai dengan data. Apabila tidak terdapat informasi apapun tentang bentuk fungsi, maka digunakan pendekatan nonparametrik (Hardle, 1990). Pendekatan nonparametrik tidak tergantung pada asumsi bentuk kurva tertentu, sehingga memberikan fleksibilitas lebih besar. Pendugaan fungsi regresi nonparametrik dilakukan berdasarkan data menggunakan teknik pemulusan seperti histogram (Green and Silverman, 1994), penduga kernel,  $k$ -Nearest Neighbour (Hardle, 1990), deret ortogonal, penduga spline, deret Fourier dan wavelet (Eubank, 1988).

Penduga polinomial lokal diperkenalkan oleh Rosenblatt (1956) dan Parzen (1962) dengan memberi bobot pada setiap pengamatan (Hardle, 1990). Pendugaan ini didasarkan pada prinsip meminimumkan jumlah kuadrat sisaan dengan pembobot fungsi kernel, sedangkan ukuran bobot ditentukan oleh parameter *bandwidth*. Polinomial lokal merupakan suatu pendekatan yang fleksibel dan efisien dalam metode statistika karena memiliki orde

sesuai dengan yang diinginkan. Orde ini yang akan menjadi derajat polinomial lokal yang sesuai bagi fungsi regresi (Fan and Gijbels, 1996). Beberapa fungsi kernel yang biasa digunakan dalam menduga model regresi nonparametrik menggunakan pendekatan polinomial lokal, antara lain *uniform*, *triangle*, Epanechnikov, Gaussian, kuartik dan cosinus.

Dalam penelitian ini akan dibahas pendugaan model regresi nonparametrik menggunakan pendekatan fungsi kernel Gaussian pada pasien penderita penyakit Tuberkulosis Paru rumah sakit Syaiful Anwar Malang.

### **1.2 Rumusan Masalah**

Rumusan masalah dalam penelitian ini adalah:

1. Bagaimana mendapatkan penduga fungsi  $m(X_i)$  pada regresi nonparametrik dengan penduga kernel?
2. Bagaimana menerapkan model regresi nonparametrik menggunakan penduga kernel pada pasien penderita penyakit Tuberkulosis Paru Rumah Sakit Dr Saiful Anwar (RSSA) Malang?

### **1.3 Tujuan Penulisan**

Tujuan dari penelitian ini adalah:

1. Mendapatkan penduga fungsi  $m(X_i)$  pada regresi nonparametrik dengan penduga kernel.
2. Menerapkan model regresi nonparametrik menggunakan penduga kernel pada pasien penderita penyakit Tuberkulosis Paru Rumah Sakit Dr Saiful Anwar (RSSA) Malang.

### **1.4 Manfaat Penelitian**

Manfaat yang diharapkan dari hasil penelitian ini :

1. Mengkaji model regresi nonparametrik menggunakan penduga kernel.
2. Mendapatkan penduga model regresi nonparametrik dengan pendekatan kernel pada pasien penderita penyakit Tuberkulosis Paru Rumah Sakit Dr Saiful Anwar (RSSA) Malang.



### 1.5 Batasan Masalah

Mengacu pada rumusan masalah penelitian ini dibatasi pada pendekatan polinomial lokal ber-orde satu menggunakan fungsi kernel Gaussian pada data bivariat.

UNIVERSITAS BRAWIJAYA



UNIVERSITAS BRAWIJAYA



## BAB II TINJAUAN PUSTAKA

### 2.1 Analisis Regresi Parametrik

Analisis regresi parametrik menggambarkan hubungan antara peubah prediktor dengan peubah respon untuk menduga pola hubungan antar peubah (Eubank, 1988). Dalam pendekatan parametrik terdapat asumsi mengenai model regresi yang dapat diwakili oleh suatu model parametrik, misal kurva berbentuk linier, kuadratik, kubik dan polinomial derajat  $p$ .

Bentuk hubungan antara peubah respon dengan satu peubah prediktor secara matematis dituliskan dalam model regresi:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad i = 1, 2, \dots, n \quad (2.1)$$

di mana :

$n$  = banyaknya pengamatan

$\beta_0$  = *intersep*

$\beta_1$  = koefisien regresi peubah prediktor

$\varepsilon_i$  = sisaan ke- $i$

Sebelum melakukan pendugaan parameter  $\beta_0$  dan  $\beta_j$ , terdapat asumsi yang harus dipenuhi agar didapatkan penduga yang baik yaitu sisaan menyebar normal dengan rata-rata 0, ragam  $\sigma^2$  dan antar sisaan saling bebas.

$$\varepsilon_i \sim \text{NIID}(0, \sigma^2)$$

$$E(\varepsilon_i) = 0$$

$$V(\varepsilon_i) = \sigma^2$$

$$\text{cov}(\varepsilon_i, \varepsilon_j) = 0$$

$$E(Y_i) = E(\beta_0 + \beta_1 X_i + \varepsilon_i)$$

$$= E(\beta_0 + \beta_1 X_i) + E(\varepsilon_i)$$

$$= E(\beta_0 + \beta_1 X_i) + 0$$

$$= E(\beta_0 + \beta_1 X_i)$$

$$\hat{Y}_i = \beta_0 + \beta_1 X_i$$

$$\varepsilon_i = Y_i - \hat{Y}_i$$

$$\varepsilon_i = Y_i - (\beta_0 + \beta_1 X_i)$$

Pendugaan parameter  $\beta_0$  dan  $\beta_1$  dilakukan menurut metode kuadrat terkecil (MKT) melalui pememinimuman jumlah kuadrat sisaan, dengan menurunkan jumlah kuadrat sisaan secara parsial terhadap  $\beta_0$  dan  $\beta_1$  kemudian disamakan dengan nol menghasilkan persamaan:

$$JKG = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2 \quad (2.2)$$

$$\begin{aligned} \frac{\partial JKG}{\partial \beta_0} &= \frac{\partial \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2}{\partial \beta_0} \\ &= 2 \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)(-1) \\ 0 &= -2 \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) \\ 0 &= \sum_{i=1}^n Y_i - \sum_{i=1}^n \hat{\beta}_0 - \sum_{i=1}^n \hat{\beta}_1 X_i \\ 0 &= \sum_{i=1}^n Y_i - n \hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^n X_i \\ n \hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^n X_i &= \sum_{i=1}^n Y_i \quad (2.3) \\ \frac{\partial JKG}{\partial \beta_1} &= \frac{\partial \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2}{\partial \beta_1} \\ 0 &= 2 \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)(-X_i) \\ 0 &= \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n \hat{\beta}_0 X_i - \sum_{i=1}^n \hat{\beta}_1 X_i^2 \\ 0 &= \sum_{i=1}^n X_i Y_i - \hat{\beta}_0 \sum_{i=1}^n X_i - \hat{\beta}_1 \sum_{i=1}^n X_i^2 \end{aligned}$$

$$\hat{\beta}_0 \sum_{i=1}^n X_i + \hat{\beta}_1 \sum_{i=1}^n X_i^2 = \sum_{i=1}^n X_i Y_i \quad (2.4)$$

Persamaan normal (2.3) dan (2.4) berbentuk matriks:

$$\begin{bmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i Y_i \end{bmatrix}$$

$$(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{Y}$$

di mana  $\mathbf{Y}_{n \times 1} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$ ,  $\mathbf{X}_{n \times 2} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}$ , dan  $\hat{\boldsymbol{\beta}}_{2 \times 1} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix}$

Setiap ruas persamaan normal dikalikan dengan  $(\mathbf{X}'\mathbf{X})^{-1}$ :

$$(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

karena  $(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X}) = \mathbf{I}$ , sehingga penduga bagi  $\boldsymbol{\beta}$  adalah:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \quad (2.5)$$

(Draper dan Smith, 1992)

## 2.2 Regresi Nonparametrik

Eubank (1988) menyatakan bahwa regresi nonparametrik digunakan untuk mengetahui hubungan antara peubah respon dengan peubah prediktor ketika bentuk hubungan antara kedua peubah tidak diketahui dan tidak diketahui informasi sebelumnya. Misalkan terdapat peubah respon ( $Y$ ) dan satu peubah prediktor ( $X$ ), maka hubungan antara peubah dapat dinyatakan dalam model regresi nonparametrik berikut :

$$Y_i = m(X_i) + \varepsilon_i \quad i=1,2,\dots,n \quad (2.6)$$

di mana:

- $Y_i$  = nilai pengamatan ke- $i$  peubah respon
- $m(X_i)$  = fungsi regresi yang ingin diduga, namun tidak diketahui bentuknya
- $X_i$  = nilai ke- $i$  peubah prediktor
- $\varepsilon_i$  = sisaan ke- $i$  yang menyebar normal dengan rata-rata 0 dan ragam  $\sigma^2$

Pendekatan regresi nonparametrik memiliki fleksibilitas tinggi, karena tidak memerlukan informasi mengenai fungsi regresi dan  $m(X_i)$  akan mengikuti bentuk data, dengan kata lain data diharapkan mencari sendiri bentuk penduga kurva regresi tanpa dipengaruhi oleh faktor subyektifitas peneliti (Eubank, 1988).

Pendugaan fungsi regresi nonparametrik  $m(X_i)$  dilakukan berdasarkan data menggunakan teknik pemulusan tertentu.

### 2.2.1 Fungsi Kernel

Penduga kernel diperkenalkan oleh Rosenblatt (1956) dan Parzen (1962). Secara umum fungsi kernel ( $K_h$ ) dengan pembobot *bandwidth* didefinisikan sebagai:

$$K_h(u) = \frac{1}{h} K\left(\frac{u}{h}\right), \text{ untuk } -\infty < u < \infty \text{ dan } h > 0 \quad (2.7)$$

dan memenuhi sifat:

(i)  $K_h(u) \geq 0$ , untuk semua  $x$

(ii)  $\int_{-\infty}^{\infty} K_h(u) du = 1$

(iii)  $\int_{-\infty}^{\infty} x^2 K_h(u) du = \sigma^2 > 0$

(iv)  $\int_{-\infty}^{\infty} x K_h(u) du = 0$

Penduga kernel untuk fungsi kepekatan  $f_h(u)$  adalah :

$$f_h(u) = \frac{1}{n} \sum_{i=1}^n K_h(X_i - X) = \frac{1}{n_h} \sum_{i=1}^n K\left(\frac{X_i - X}{h}\right) \quad (2.8)$$

Dari persamaan (2.7) terlihat bahwa  $\hat{f}_h(u)$  tergantung pada fungsi kernel  $K_h$  dan parameter pemulus  $h$ . Bentuk bobot kernel ditentukan oleh fungsi kernel  $K_h$ , sedangkan ukuran bobot ditentukan oleh parameter pemulus  $h$  yang disebut *bandwidth*.



Beberapa macam fungsi kernel yang sering digunakan sebagai bobot adalah:

1. Kernel Uniform :  $K_h(u) = \frac{1}{2}, -1 < u < 1$
2. Kernel Triangle :  $K_h(u) = 1 - |u|, -1 < u < 1$
3. Kernel Gaussian :  $K_h(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right), -\infty < u < \infty$

di mana  $u_i = \frac{X_i - X}{h}, i = 1, 2, 3, \dots, n$

$X$  adalah peubah prediktor,  $X_i$  adalah nilai ke- $i$  peubah prediktor, dan  $h$  adalah *bandwith* (Hardle, 1990).

### 2.2.2 Regresi Polinomial Lokal

Regresi polinomial lokal adalah regresi nonparametrik di mana fungsi regresi  $m(X_i)$  berbentuk polinomial. Pada regresi polinomial biasa, kurva regresi diplot pada seluruh wilayah data, sedangkan dalam regresi polinomial lokal persamaan diplot hanya dalam lingkungan  $X_i$  saja. Hal ini disebabkan oleh fakta bahwa jika hanya satu polinomial saja yang diplot untuk seluruh wilayah data dengan kondisi kompleks, maka kemungkinan persamaan tersebut cocok untuk wilayah data tertentu namun tidak cocok untuk wilayah data lain. Pertimbangan-pertimbangan ini yang mendasari konsep pembagian data ke dalam beberapa wilayah dan memplotkan kurva untuk masing-masing wilayah (Nugraha, 2009).

### 2.2.3 Metode Kuadrat Terkecil Terboboti

Regresi kernel adalah teknik statistika nonparametrik yang digunakan untuk menduga harapan bersyarat dari peubah respon ( $Y$ ) relatif terhadap peubah prediktor ( $X$ ) yang dapat ditulis  $E(Y/X) = m(X_i)$  atau  $\hat{Y} = m(X_i)$  dengan  $E(Y/X = x) = \int y \frac{f(x,y)}{f(x)} dy$  di mana  $m(X_i)$  adalah fungsi yang tidak diketahui dan untuk mendapatkannya digunakan bobot yang sesuai. Pendekatan kernel dapat digunakan menduga kurva regresi  $m(X_i)$  pada regresi nonparametrik. Untuk mendapatkan penduga kernel digunakan polinomial lokal (Wand and Jones, 1995).

Pertama kurva regresi  $m(X_i)$  dihipotesiskan oleh polinomial berikut:

$$m(X_i) = \beta_0 + \beta_1(X_1 - X)^1 + \dots + \beta_p(X_n - X)^p \quad (2.9)$$

$p$  adalah derajat polinomial. Penduga  $\beta$  dilandasi pada metode kuadrat terkecil terboboti dengan meminimumkan :

$$\sum_{i=1}^n \varepsilon^2 K_h(X_i - X) = \sum_{i=1}^n (Y_i - \beta_0 - \dots - \beta_p(X_i - X)^p)^2 (K_h(X_i - X)) \quad (2.10)$$

di mana  $K_h(X_i - X)$  merupakan pembobot kernel, persamaan (2.10) disajikan dalam bentuk matriks:

$$\mathbf{Q} = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{W} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \quad (2.11)$$

di mana

$$\mathbf{Y}'_{1 \times n} = (Y_1, Y_2, \dots, Y_n)$$

$$\boldsymbol{\beta}'_{1 \times (p+1)} = (\beta_0, \beta_1, \beta_2, \dots, \beta_n)$$

$$\mathbf{X}_{n \times (p+1)} = \begin{bmatrix} 1 & (X_1 - X) & \dots & (X_1 - X)^p \\ \vdots & \vdots & & \vdots \\ 1 & (X_n - X) & \dots & (X_n - X)^p \end{bmatrix}$$

$$\mathbf{W}_{n \times n} = \begin{bmatrix} K_h(X_1 - X) & 0 & \dots & 0 \\ 0 & \vdots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & K_h(X_n - X) \end{bmatrix} \quad (2.12)$$

Persamaan (2.12) dapat diuraikan menjadi:

$$\mathbf{Q} = (\mathbf{Y}' - \boldsymbol{\beta}' \mathbf{X}') (\mathbf{W} \mathbf{Y} - \mathbf{W} \mathbf{X} \boldsymbol{\beta})$$

$$\mathbf{Q} = \mathbf{Y}' \mathbf{W} \mathbf{Y} - \mathbf{Y}' \mathbf{W} \mathbf{X} \boldsymbol{\beta} - \boldsymbol{\beta}' \mathbf{X}' \mathbf{W} \mathbf{Y} + \boldsymbol{\beta}' \mathbf{X}' \mathbf{W} \mathbf{X} \boldsymbol{\beta}$$

karena  $\mathbf{Y}' \mathbf{W} \mathbf{X} \boldsymbol{\beta}$  adalah skalar maka

$$(\mathbf{Y}' \mathbf{W} \mathbf{X} \boldsymbol{\beta})' = \boldsymbol{\beta}' \mathbf{X}' \mathbf{W} \mathbf{Y} \text{ sehingga}$$

$$\mathbf{Q} = \mathbf{Y}' \mathbf{W} \mathbf{Y} - \boldsymbol{\beta}' \mathbf{X}' \mathbf{W} \mathbf{Y} - \boldsymbol{\beta}' \mathbf{X}' \mathbf{W} \mathbf{Y} + \boldsymbol{\beta}' \mathbf{X}' \mathbf{W} \mathbf{X} \boldsymbol{\beta}$$

$$\mathbf{Q} = \mathbf{Y}' \mathbf{W} \mathbf{Y} - 2\boldsymbol{\beta}' \mathbf{X}' \mathbf{W} \mathbf{Y} + \boldsymbol{\beta}' \mathbf{X}' \mathbf{W} \mathbf{X} \boldsymbol{\beta} \quad (2.13)$$

Penduga  $\hat{\beta}$  diperoleh dengan menurunkan persamaan (2.13) terhadap  $\beta$ . Nilai minimum  $Q$  diperoleh pada saat  $\frac{\partial Q}{\partial \beta} = 0$ , sehingga menghasilkan:

$$\begin{aligned}\frac{\partial Q}{\partial \beta} &= -2X'WY + 2X'WX\beta = 0 \\ 0 &= -2(X'WY + X'WX\beta) \\ 0 &= (X'WY + X'WX\beta)\end{aligned}$$

persamaan ini memberikan :

$$\begin{aligned}X'WX\beta &= X'WY, \text{ atau} \\ (X'WX)^{-1}X'WX\beta &= (X'WX)^{-1}X'WY\end{aligned}$$

sehingga diperoleh

$$\hat{\beta} = (X'WX)^{-1}X'WY \quad (2.14)$$

#### 2.2.4 Pemilihan *Bandwidth* Optimal

*Bandwidth* ( $h$ ) adalah parameter pemulus (*smoothing*) yang digunakan untuk mengontrol kemulusan dari kurva yang diduga dan sebagai ukuran kesesuaian fungsi pada data.

Permasalahan utama pada *smoothing* kernel bukan terletak pada pemilihan fungsi kernel, tetapi pada pemilihan *bandwidth*. Oleh karena itu perlu dilakukan pemilihan *bandwidth* optimal. Suatu kriteria untuk *bandwidth* akan dibatasi pada kelas penduga linier, untuk semua  $h$  terdapat matriks  $A(h)$  berukuran  $n \times n$  sehingga:

$$\hat{m}_h(X_i) = A(h)Y = X\hat{\beta} \quad (2.15)$$

$$A(h)Y = X(X'WX)^{-1}X'WY$$

$$A(h)YY' = X(X'WX)^{-1}X'WYY'$$

$$A(h)YY'(YY')^{-1} = X(X'WX)^{-1}X'WYY'(YY')^{-1}$$

$$A(h) = X(X'WX)^{-1}X'W \quad (2.16)$$

Pemilihan *bandwidth* optimal ditekankan pada penyeimbangan antara bias dan ragam. Satu perumusan masalah yang dapat memperlihatkan hubungan antara bias dan ragam adalah *Mean Square Error (MSE)*.  $MSE(\hat{m}_h(X))$  merupakan alat ukur yang tepat

untuk mengukur *goodness of fit* suatu model terhadap parameter pemulus (bandwidth).  $MSE(\hat{m}_h(X))$  didefinisikan sebagai:

$$MSE(\hat{m}_h(X_i)) = \frac{1}{n} \sum_{i=1}^n e_i^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (2.17)$$

Persamaan (2.17) dapat disajikan dalam bentuk:

$$MSE(\hat{m}_h(X_i)) = E\{\hat{m}_h(X_i) - m_h(X_i)\}^2 \quad (2.18)$$

dan diuraikan menjadi:

$$\begin{aligned} MSE(\hat{m}_h(X_i)) &= E\{\hat{m}_h(X_i) - m_h(X_i)\}^2 \\ &= E\left[\left(\hat{m}_h(X) - E(\hat{m}_h(X))\right) + E(\hat{m}_h(X)) - m_h(X)\right]^2 \\ &= E\left[\hat{m}_h(X_i) - E(\hat{m}_h(X_i))\right]^2 + \left[E(\hat{m}_h(X_i)) - m_h(X_i)\right]^2 \\ &\quad + 2E\left[\left(\hat{m}_h(X_i) - E(\hat{m}_h(X_i))\right)\left(E(\hat{m}_h(X_i)) - m_h(X_i)\right)\right] \\ &= E\left[\hat{m}_h(X_i) - E(\hat{m}_h(X_i))\right]^2 + \left[E(\hat{m}_h(X_i)) - m_h(X_i)\right]^2 \\ &= \text{Var}(\hat{m}_h(X_i)) + \left[\text{Bias}(\hat{m}_h(X_i))\right]^2 \end{aligned} \quad (2.19)$$

di mana  $\text{Var}(\hat{m}_h(X_i)) = E\left[\hat{m}_h(X_i) - E(\hat{m}_h(X_i))\right]^2$  dan

$$\text{Bias}(\hat{m}_h(X_i)) = E\left[\hat{m}_h(X_i) - m_h(X_i)\right]$$

Ragam pada persamaan (2.19) akan mendekati proporsional pada saat  $\frac{1}{nh}$ , sehingga  $h$  dipilih yang besar. Sedangkan bias pada persamaan (2.19) mendekati proporsional pada saat  $h$  kuadrat, sehingga  $h$  dipilih yang kecil. Dapat juga dilihat dari definisi berikut:

$$\text{Var}(\hat{m}_h(X_i)) \approx \frac{1}{nh} \text{ dan } \text{Bias}(\hat{m}_h(X_i)) \approx h^2$$

Hal ini berarti :

1. Nilai *bandwidth* kecil akan menghasilkan bias kecil tetapi ragam yang besar.
2. Nilai *bandwidth* besar akan menghasilkan bias besar tetapi ragam kecil.

Salah satu metode untuk mendapatkan *bandwidth* optimal adalah dengan menggunakan kriteria *Generalized Cross Validation* (GCV) :

$$GCV(h) = \frac{MSE(\hat{m}_n(X_i))}{\left(\frac{1}{n} \text{tr}(I - A(h))\right)^2} \quad (2.20)$$

*Bandwidth* optimal ( $h_{opt}$ ) diperoleh dari *bandwidth* yang menghasilkan GCV minimum dan dapat dirumuskan sebagai:

$$h_{opt} = \min GCV(h) \quad (2.21)$$

### 2.3 Koefisien Determinasi

Koefisien determinasi ( $R^2$ ) merupakan besaran yang digunakan untuk mengukur kelayakan model regresi dan menunjukkan besar kontribusi  $X$  terhadap perubahan  $Y$ . Semakin tinggi nilai  $R^2$  (mendekati 1) semakin baik model regresi yang terbentuk :

$$R^2 = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} = \frac{SSR}{SST} \quad (2.22)$$

di mana:

- $\hat{Y}_i$  = pendugan nilai ke- $i$  peubah respon
- $\bar{Y}$  = rata-rata peubah respon
- $Y_i$  = nilai ke- $i$  peubah respon
- SSR = jumlah kuadrat regresi
- SST = jumlah kuadrat total

Menurut Draper dan Smith (1992), ada statistik lain sebagai fungsi  $R^2$ , yaitu koefisien determinasi yang disesuaikan  $R^2_{adjusted}$

$$R^2_{adjusted} = 1 - \frac{SSE/(n-p)}{SST/(n-1)} = 1 - (1 - R^2) \left( \frac{n-1}{n-p} \right) \quad (2.23)$$

di mana :

- SSE = Jumlah Kuadrat Sisaan
- SST = Jumlah Kuadrat Total
- $p$  = Banyaknya parameter model
- $n$  = Banyaknya pengamatan



Pada  $R^2_{adjusted}$  telah dilakukan penyesuaian terhadap derajat bebas galat dan total. Penyesuaian dilakukan agar  $R^2_{adjusted}$  dapat digunakan tidak hanya untuk membandingkan beberapa persamaan regresi yang diterapkan pada segugus data, tetapi juga untuk membandingkan persamaan regresi dari dua atau lebih gugus data (Draper dan Smith, 1992).

## 2.4 Tuberkulosis (TB)

Penyakit Tuberkulosis paru telah dikenal lebih dari satu abad yang lalu, yakni sejak ditemukannya kuman penyebab Tuberkulosis oleh Robert Koch tahun 1882, namun sampai saat ini penyakit Tuberkulosis (TB) masih tetap menjadi permasalahan kesehatan di seluruh dunia dan sebagai penyebab kematian utama yang diakibatkan oleh penyakit infeksi. Pada April 1993 WHO menyatakan TB sebagai suatu masalah kesehatan masyarakat yang sangat serius di seluruh dunia, karena satu dari tiga penduduk dunia diperkirakan telah terinfeksi oleh bakteri *Mycobacterium tuberculosis* (disebut juga Basil Tahan Asam = BTA) sebagai kuman penyebab TB yang dibuktikan dengan pemeriksaan Mantoux tes. Sekitar 95% penderita TB terdapat di negara sedang berkembang dengan sosioekonomi rendah termasuk Indonesia dan 75% dari penderita TB terjadi pada usia produktif. Setiap tahun terdapat sekitar 4 juta penderita baru TB paru menular di dunia. Menurut WHO diperkirakan pada tahun 2000 terdapat 8,74 juta penderita baru TB dan akan menjadi 10,2 juta penderita baru TB pada tahun 2005. Di kawasan Asia Tenggara diduga terjadi lebih dari 3,5 juta penderita baru TB dan lebih dari 1,3 juta kematian akibat penyakit ini, dan diperkirakan pada tahun 2005 terdapat 3 juta penderita baru TB.

Infeksi TB paru diawali oleh inhalasi droplet yang mengandung bakteri *M. tuberculosis*. Di dalam paru bakteri akan difagosit oleh makrofag alveolar dan akan menginduksi respon proinflamasi lokal yang selanjutnya akan terjadi rekrutmen sel mononuklear dari pembuluh darah di sekitarnya. Sel-sel tersebut akan mulai membentuk *granuloma* atau tuberkel, yang merupakan bentukan khas dari tuberkulosis. *Granuloma* terdiri dari makrofag yang terinfeksi, dikelilingi oleh *foamy macrophage* dan sel mononuklear yang lainnya, diselimuti oleh limfosit yang



berhubungan dengan jaringan fibrosis dan komponen matriks ekstraseluler yang lainnya. Proses pembentukan *granuloma* ini adalah untuk melindungi, agar tidak terjadi penyebaran ke jaringan yang lainnya. Pada perjalanan berikutnya *granuloma* akan semakin berkembang, dimana jaringan fibrosis semakin banyak menyelimuti dan pembuluh darah akan semakin berkurang. Keadaan ini akan menyebabkan nekrosis sentral akibat hipoksemia (Russel, 2007).

Hasil pemeriksaan plasma pada penderita tuberkulosis paru aktif didapatkan kadar suPAR yang lebih tinggi dibandingkan individu normal. Peningkatan kadar suPAR ini mencerminkan adanya mobilisasi makrofag ke dalam bronkus. Dalam proses perlekatan dan migrasi sel monosit ke tempat infeksi melibatkan interaksi antara suPAR dan integrin. Digambarkan dengan jelas adanya ikatan antara uPAR dan CD11b/CD18 (*complement receptor 3 atau MAC1*) yang merupakan interaksi antara suPAR dan integrin.

Dalam penelitiannya Juffermanas (1999) menjelaskan adanya upregulasi gabungan antara uPAR dan CD11b selama endotoksemia dalam eksperimen penyuntikan antigen *lipoarabinomannan (LAM)* yang menyusun komposisi dinding sel *M. tuberculosis*. Hal yang penting adalah CD11b/CD1 juga memediasi perlekatan *M. tuberculosis* pada makrofag, keadaan ini menunjukkan bahwa tuberkulosis memanfaatkan upregulasi perlekatan tersebut dan mempengaruhi rekrutmen makrofag (Oelsen *et al*, 2002).



UNIVERSITAS BRAWIJAYA



## BAB III METODOLOGI

### 3.1 Sumber Data

Data yang digunakan diambil dari penelitian Chozin (2009) berupa kadar monocyte pasien penderita penyakit Tuberkulosis Paru. Penelitian dilakukan di Ruang Rawat Jalan Poli Paru dan Ruang Rawat Inap Rumah Sakit Dr Saiful Anwar (RSSA) Malang periode tahun 2008.

### 3.2 Metode Analisis Data

Berikut langkah-langkah pendugaan regresi nonparametrik dengan pendekatan fungsi kernel.

1. Menduga model regresi nonparametrik dengan pendekatan fungsi kernel dengan langkah-langkah sebagai berikut:
  - a. Memasukan data  $(X_i, Y_i)$  untuk  $i = 1, 2, \dots, n$  yang memenuhi model regresi nonparametrik  $Y_i = m_h(X_i) + \varepsilon_i$  dan mengurutkan data berdasarkan  $X_i$ .
  - b. Menentukan orde polinomial lokal  $p$  untuk menduga fungsi regresi dengan penduga fungsi kernelsesuai persamaan (2.9).
2. Menentukan nilai GCV untuk setiap nilai  $h$  menurut langkah berikut:
  - a. Menghitung fungsi kernel.
  - b. Menghitung matriks  $W$  sesuai persamaan (2.12).
  - c. Menentukan matriks  $A(h)$  sesuai persamaan (2.16).

$$\hat{m}(X_i) = \sum_{j=0}^p \beta_j (X_i - X)^j.$$

- d. Menghitung nilai  $MSE(h)$  sesuai persamaan (2.17).

$$MSE(\hat{m}_h(X)) = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{m}(X_i))^2$$

dengan mengubah ke dalam bentuk matriks

$$MSE(\hat{m}_h(X)) = \mathbf{n}^{-1} ((\mathbf{Y} - \mathbf{A}(h)\mathbf{Y})' (\mathbf{Y} - \mathbf{A}(h)\mathbf{Y}))$$

- e. Menghitung nilai  $GCV(h)$  sesuai persamaan (2.20).

$$GCV(h) = \frac{MSE(\hat{m}_h(X_i))}{\left(\frac{1}{n} \text{tr}(1 - A(h))\right)^2}$$

3. Mendapatkan penduga  $\beta$  yaitu  $\hat{\beta}$  yang diperoleh dengan meminimumkan jumlah kuadrat terkecil terboboti (JKTT), menurut persamaan (2.14).

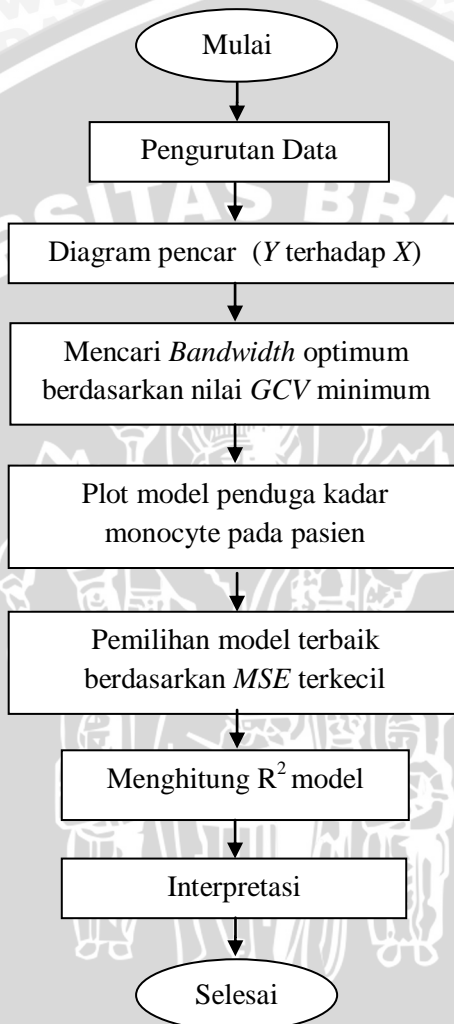
$$\hat{\beta} = (X'WX)^{-1}X'WY$$

4. Mendapatkan penduga model regresi nonparametrik dengan pendekatan Nadaraya-Watson dengan fungsi kernel:

$$\hat{Y} = \hat{m}(X_i) = X\hat{\beta}$$

5. Menghitung  $R_{adjusted}^2$ .





Gambar 3.1. Diagram Alir Regresi Nonparametrik

UNIVERSITAS BRAWIJAYA





## BAB IV HASIL DAN PEMBAHASAN

### 4.1 Penduga Kurva Regresi Nonparametrik

Diberikan data berpasangan  $(X_i, Y_i)$  yang di asumsikan mengikuti model regresi

$$Y_i = m(X_i) + \varepsilon_i \quad i=1,2,\dots,n \quad (4.1)$$

dengan  $\varepsilon_i$  adalah sisaan pengamatan ke- $i$  yang menyebar normal dengan rata-rata 0 dan ragam  $\sigma^2$ . Fungsi  $m(X_i)$  adalah fungsi regresi yang akan diduga, namun tidak diketahui bentuknya. Dalam regresi nonparametrik untuk menduga kurva regresi  $m(X_i)$  dapat digunakan pendekatan kernel. Penduga kernel yang digunakan adalah penduga Nadaraya-Watson. Untuk mendapatkan penduga kernel digunakan polinomial lokal (Wand and Jones, 1995) sehingga akan diperoleh penduga  $\hat{\beta}$  (persamaan 2.14) sebagai berikut:

$$\hat{\beta} = (X'WX)^{-1}X'WY$$

Penduga  $m(x_i)$  diberikan dalam bentuk :

$$\hat{m}(x, p, h) = \mathbf{e}'_1(X'WX)^{-1}X'WY \quad (4.2)$$

di mana  $\mathbf{e}_1$  adalah vektor  $(p+1) \times 1$ . Jika  $p=0$  akan diperoleh penduga Nadaraya-Watson (Wand and Jones, 1995).

Untuk  $p=0$  maka :

$$\hat{m}(x,0, h) = \mathbf{e}'_1(X'WX)^{-1}X'WY,$$

Dengan  $\mathbf{e}'_1 = 1$

$$Y'_{1 \times n} = (Y_1, Y_2, \dots, Y_n)$$

$$X_{n \times p} = \begin{bmatrix} 1 & X_1 - X & \dots & (X_1 - X)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_n - X & \dots & (X_n - X)^p \end{bmatrix}$$

$$W_{n \times n} = \begin{bmatrix} K_h(X_1 - X) & 0 & \dots & 0 \\ 0 & & \ddots & \vdots \\ \vdots & & & 0 \\ 0 & \dots & & K_h(X_n - X) \end{bmatrix}$$

maka

$$\hat{m}_h(x, 0, h) = \hat{\beta} \begin{bmatrix} [1 \dots 1] \\ \begin{bmatrix} K_h(X_1 - X) & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & & 0 \\ 0 & \dots & K_h(X_n - X) \end{bmatrix} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}^{-1} \cdot A$$

dengan

$$A = \begin{bmatrix} [1 \dots 1] \\ \begin{bmatrix} K_h(X_1 - X) & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & & 0 \\ 0 & \dots & K_h(X_n - X) \end{bmatrix} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

$$= \begin{bmatrix} [K_h(X_1 - X) & K_h(X_2 - X) & \dots & K_h(X_n - X)] \\ \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \end{bmatrix}^{-1} \cdot B$$

dengan

$$B = \begin{bmatrix} [K_h(X_1 - X) & K_h(X_2 - X) & \dots & K_h(X_n - X)] \\ \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \end{bmatrix}$$

$$= [K_h(X_1 - x) + K_h(X_2 - X) + \dots + K_h(X_n - X)]^{-1} \cdot C$$

dengan

$$C = [K_h(X_1 - X)Y_1 + K_h(X_2 - X)Y_2 + \dots + K_h(X_n - X)Y_n]$$

$$= \left[ \frac{1}{\sum_{i=1}^n K_h(X_i - X)} \right] \left[ \sum_{i=1}^n K_h(X_i - X)Y_i \right]$$

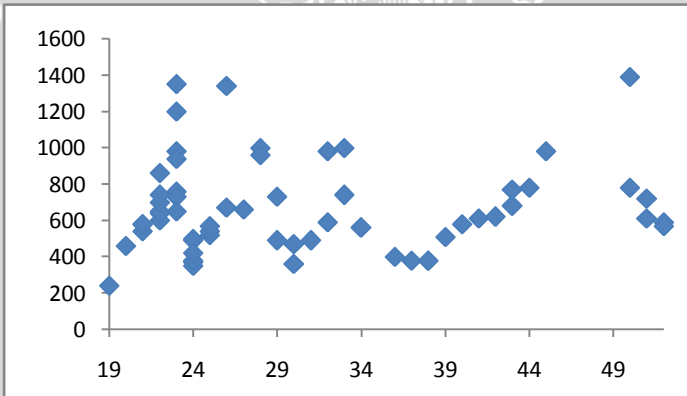
sehingga diperoleh penduga :

$$\hat{m}(X_i) = \frac{\sum_{i=1}^n K_h(X_i - X) Y_i}{\sum_{i=1}^n K_h(X_i - X)} \quad (4.3)$$

Penduga ini merupakan penduga kernel Nadaraya-Watson.

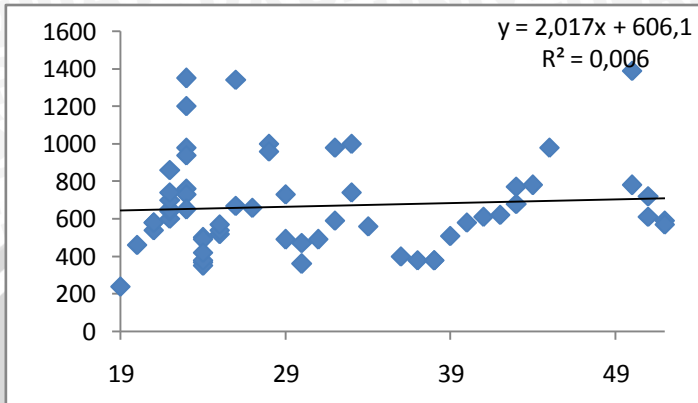
#### 4.2 Penduga Kernel

Data yang digunakan adalah data kadar monocyte pasien penderita penyakit tuberkulosis, dengan banyak data  $n=60$ , dengan peubah prediktor adalah usia pasien dan peubah respon adalah kadar monocyte. Disajikan dalam diagram pencar kedua peubah (Gambar 4.1).

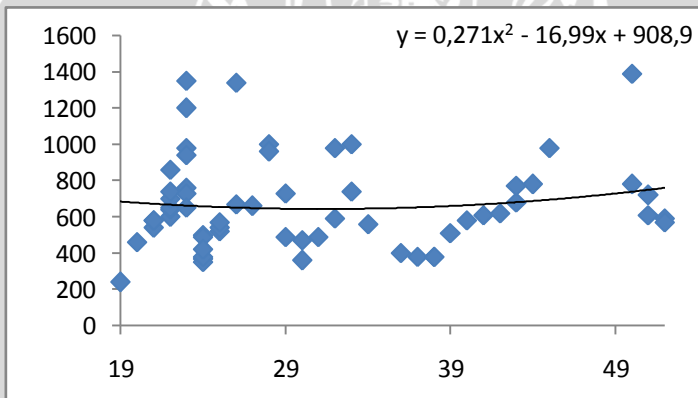


Gambar 4.1 Plot Kadar Monocyte dengan usia pasien (tahun)

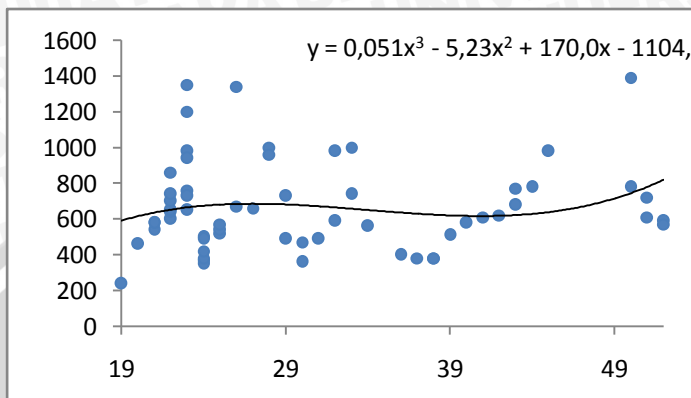
Gambar 4.1 menunjukkan bentuk pola antara kadar monocyte dengan usia tidak memperlihatkan bentuk pola tertentu. Dari plot data antara kedua peubah akan diduga menggunakan pendekatan regresi parametrik, berikut merupakan hasil pendugaan menggunakan regresi parametrik:



Gambar 4.2 Model regresi linier kadar monocyte dengan usia pasien (tahun)



Gambar 4.3 Model regresi kuadratik kadar monocyte dengan usia pasien (tahun)



Gambar 4.4 Model regresi kubik kadar monocyte dengan usia pasien (tahun)

Dari hasil pendugaan di atas diketahui jika pendekatan regresi parametrik tidak cocok untuk digunakan dalam memodelkan data kadar monocyte penderita penyakit tuberkulosis, hal ini dapat dilihat dari nilai *R Square* ( $R^2$ ) yang sangat kecil. Menurut Draper dan Smith (1992), *R Square* ( $R^2$ ) digunakan untuk mengukur kebaikan dari suatu persamaan regresi (goodness of fit), yaitu memberikan proporsi atau persentase variasi total dalam peubah respon yang dijelaskan oleh peubah prediktor. Nilai  $R^2$  terletak antara 0 – 1, dan kecocokan model dikatakan lebih baik jika  $R^2$  semakin mendekati 1. Untuk nilai  $R^2$  dapat dilihat pada table berikut:

Tabel 4.1 Nilai Koefisien Determinasi Model

Respon	Orde Polinom	$R^2$
Y1	1	0,006
	2	0,024
	3	0,152

Gambar 4.1 menunjukkan bentuk kurva yang menggambarkan hubungan antara kadar monocyte dalam darah dengan usia pasien penderita penyakit tuberkulosis. Pendekatan parametrik tidak tepat untuk digunakan pada data tersebut, karena kurva tidak membentuk pola

linier, kuadrat, eksponensial, atau kubik. Kurva regresi  $m(X_i)$  akan diduga menggunakan pendekatan nonparametrik dengan penduga kernel Nadaraya-Watson (persamaan 4.3) sebagai berikut:

$$\hat{m}(X_i) = \frac{\sum_{i=1}^n K_h(X_i - X) Y_i}{\sum_{i=1}^n K_h(X_i - X)}$$

Langkah yang dilakukan dalam menduga kurva regresi nonparametrik yaitu dengan menentukan fungsi bobot atau fungsi kernel yang akan digunakan dan ukuran bobot yang optimal (*bandwidth*). Pemilihan *bandwidth* yang optimal akan menghasilkan kurva regresi yang mulus sesuai dengan pola data, ditentukan berdasarkan kriteria nilai GCV yang minimum. Dengan metode GCV (persamaan 2.22) sebagai berikut:

$$GCV(h) = \frac{MSE(\hat{m}_h(X_i))}{\left(\frac{1}{n} \text{tr}(1 - A(h))\right)^2}$$

dengan  $A(h)$  diperoleh dari persamaan 2.15:

$$\hat{m}_h(X_i) = A(h)Y = X\hat{\beta}$$

Berdasarkan hasil perhitungan dengan *software* S-Plus, nilai *bandwidth* optimum diperoleh pada saat nilai GCV minimum sebesar 1.91373e-06 pada kenaikan  $h=0.05$  dan selang  $h$  antara 0.005 sampai 5 (Lampiran 4). Namun untuk memperoleh kekakuratan yang lebih tinggi, dapat dipilih selang kenaikan  $h$  yang lebih kecil misalkan sebesar 0.005 (Lampiran 5) atau 0.001 (Lampiran 6). Dengan kenaikan tersebut juga diperoleh *bandwidth* optimal sebesar 0.005.

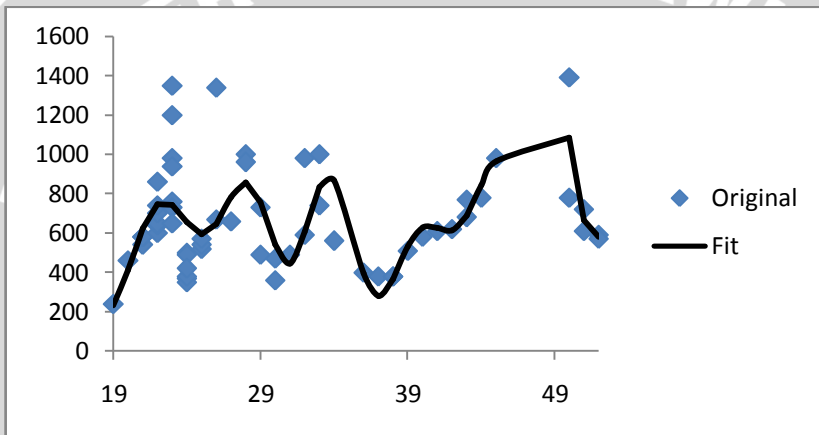
### 4.3 Pemilihan Model Regresi Nonparametrik

Berdasarkan pada hasil perhitungan dengan menggunakan *software* S-Plus, dapat diketahui bahwa fungsi kernel gaussian memiliki nilai MSE terkecil sebesar 0.827364 dengan nilai *bandwidth* sebesar 0.005, sehingga diperoleh persamaan regresi sebagai berikut:



$$\hat{m}(X_i) = \frac{\sum_{i=1}^{59} 5.643 \exp - \frac{1}{2} \left( \frac{X_i - X}{0.005} \right)^2 Y_i}{\sum_{i=1}^{59} 5.643 \exp - \frac{1}{2} \left( \frac{X_i - X}{0.005} \right)^2}$$

Untuk plot model regresi nonparametrik kadar monocyte pasien tuberkulosis paru menggunakan fungsi gaussian dengan bandwidth 0.005 dapat dilihat pada gambar 4.5.



Gambar 4.5 Model regresi nonparametrik dengan fungsi kernel gaussian

#### 4.4 Perbandingan Regresi Parametrik dengan Regresi Nonparametrik

Regresi parametrik dan regresi nonparametrik memiliki perbedaan pada pendugaan parameter. Pendugaan parameter pada regresi parametrik dilakukan menurut metode kuadrat terkecil (MKT) melalui meminimuman jumlah kuadrat sisaan, sedangkan pada regresi nonparametrik dilakukan menurut metode kuadrat terkecil terboboti (MKTT) melalui meminimuman jumlah kuadrat sisaan yang telah diberikan bobot kernel. Dalam penelitian ini perbandingan antara regresi parametrik dan regresi nonparametrik didasari pada hasil perhitungan koefisien determinasi. Hasil perhitungan koefisien determinasi dapat dilihat pada Tabel 2.

Tabel. 2. Perbandingan regresi parametrik dan regresi nonparametrik

Metode	Koefisien determinasi
Regresi Linier	0.006
Regresi Kuadratik	0.024
Regresi Kubik	0.152
Regresi Nonparametrik	0.948

Berdasarkan tabel 2, hasil perhitungan nilai koefisien determinasi regresi linier sebesar 0.006, hasil perhitungan koefisien determinasi regresi kuadratik sebesar 0.024, hasil perhitungan koefisien determinasi regresi kubik sebesar 0.152, dan hasil perhitungan koefisien determinasi regresi nonparametrik sebesar 0.948. Sesuai dengan teori koefisien determinasi maka yang dipilih nilai koefisien determinasi yang terbesar. Dengan demikian dapat disimpulkan bahwa metode regresi nonparametrik mampu menghasilkan model yang lebih baik pada data kadar monocyte pasien penderita Tuberkulosis Paru.



## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Kesimpulan yang dapat diperoleh dari penelitian ini adalah:

1. Data yang tidak berpola dan menyebar tidak beraturan, tetap dapat dicari kurva regresinya, salah satunya dengan menggunakan metode regresi nonparametrik dengan penduga kernel Nadaraya-Watson.
2. Pada data pasien penderita penyakit Tuberkulosis Paru, diperoleh *bandwidth* optimal sebesar 0.005 didasarkan kriteria nilai GCV minimum yaitu pada GCV sebesar  $1.91373 \times 10^{-6}$ . Dengan menggunakan *bandwidth* optimal tersebut, model nonparametrik dengan penduga kernel Nadaraya-Watson dengan fungsi kernel Gaussian mampu memodelkan hubungan antara usia pasien penderita penyakit Tuberkulosis Paru di rumah sakit Dr. Saiful Anwar (RSSA) dengan nilai MSE sebesar 0.827364.

#### 5.2 Saran

Berdasarkan hasil penelitian ini, maka terdapat beberapa saran yang direkomendasikan yaitu sebagai berikut:

1. Hasil penelitian ini dapat digunakan sebagai alternatif pemodelan regresi apabila hubungan antara peubah respon dan prediktor tidak menunjukkan pola atau struktur parametrik.
2. Pemilihan nilai *bandwidth* pada pendekatan nonparametrik sangat penting, maka disarankan pada penelitian selanjutnya menggunakan fungsi kernel yang berbeda untuk melihat seberapa tingkat kebaikan fungsi kernel dalam menduga model regresi nonparametrik.
3. Hasil penelitian ini dibatasi hanya pada data dengan satu peubah respon dan satu peubah prediktor, untuk itu metode ini dapat dikembangkan untuk data dengan peubah prediktor lebih dari satu (multiprediktor).

UNIVERSITAS BRAWIJAYA



## DAFTAR PUSTAKA

- Draper, N.R. dan Smith, H. 1992. *Analisis Regresi Terapan. Ed.2.* Gramedia Pustaka Utama. Jakarta.
- Chozin, I.N. 2009. *Studi Prospektif : Kadar Soluble Urokinase Plasminogen Activator Receptor (Supar) Selama Terapi Oat Pada Penderita Tuberkulosis Paru.* Fakultas Kedokteran. Universitas Brawijaya. Tidak Dipublikasikan
- Eubank, L. 1988. *Spline Smoothing and Nonparamatric Regression.* Marcell Dekker Inc. New York.
- Fan, J. and Gijbels, I. 1996. *Local Polynomial Modelling and its Applications.* Chapman & Hall. New York.
- Hardle, W. 1990. *Applied Nonparametric Regression.* Cambridge University Press. Boston.
- Husna. 2002. *Pemilihan Bandwidth dan Selang Kepercayaan dalam Regresi Nonparametrik.*(Tesis). Departemen Statistika Fakultas MIPA, Institut Sepuluh November. Surabaya. Tidak dipublikasikan.
- Jawetz dkk. 1995. *Mikrobiologi. Ed 5.* ECG. Jakarta.
- Juffermans dkk. 1999. *Elevated Chemokine Concentrations in Sera of Human Immunodeficiency Virus (HIV)-seropositive and HIV seronegative and HIV-seronegative Patiens with Tuberculosis: a possible role for mycobacterial lipoarabinomannan.* Infect and Immun. 1:Pg 4295-4297. Tanggal akses: 20 Agustus 2014.
- Nugraha, N. 2009. *Regresi Polinomial Lokal.* Departemen Matematika Fakultas MIPA. Universitas Indonesia. Depok.
- Oelsen dkk. 2002. *The Serum Level of Soluble Urokinase Receptor is Elevated in Tuberculosis Patients and Predicts Mortality during Treatment: A Community Study from Guinea-Bissau.* International Journal of Tuberculosis Lung Diseases 6(8): Pg 686-692 . Tanggal akses: 20 Agustus 2014.

Russell, D.G. 2007. *Nature Reviews Microbiology* 5. Pg 39-47. Tanggal akses: 20 Agustus 2014.

Schoot, S.R. 1997. *Matrix Analysis for Statistics*. John Wiley And Son Inc. New York.

Wand, M.P. and Jones, M.C. 1995. *Kernel Regression*. Chapman and Hall. London.

Zulfikar. 2005. *Bias, Variance and MSE Component of Kernel Estimation on Nonparametric Regression*. (Tesis). Departemen Statistika Fakultas MIPA, Institut Sepuluh November. Surabaya. Tidak dipublikasikan.





## LAMPIRAN

Lampiran 1. Kadar Monocyte Pasien Penderita Penyakit Tuberkulosis Paru Rumah Sakit Dr Saiful Anwar (RSSA), Malang

No	Usia Pasien (Tahun)	Kadar Monocyte	No	Usia Pasien (Tahun)	Kadar Monocyte
1	19	240	31	28	1000
2	20	460	32	28	960
3	21	540	33	29	730
4	21	580	34	29	490
5	21	580	35	30	360
6	22	600	36	30	470
7	22	640	37	31	490
8	22	650	38	32	590
9	22	700	39	32	980
10	22	740	40	33	1000
11	22	860	41	33	740
12	23	1200	42	34	560
13	23	1350	43	36	400
14	23	980	44	37	380
15	23	940	45	38	380
16	23	760	46	38	380
17	23	730	47	39	510
18	23	650	48	40	580
19	24	500	49	41	610
20	24	490	50	42	620
21	24	370	51	43	680
22	24	350	52	43	770
23	24	380	53	44	780
24	24	420	54	45	980
25	25	520	55	50	1390
26	25	540	56	50	780
27	25	570	57	51	720
28	26	670	58	51	610
29	26	1340	59	52	590
30	27	660	60	52	570

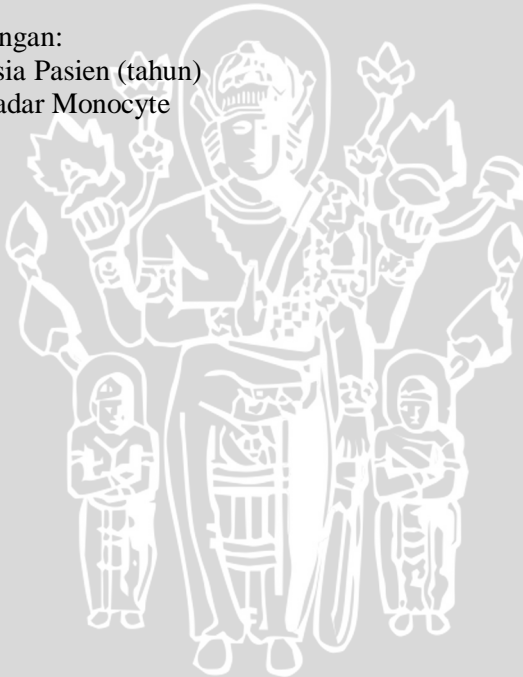
Lampiran 2. Statistika Deskriptif Kadar Monocyte Pasien Penderita Penyakit Tuberkulosis Paru

	X	Y
Min	: 19.0000	240.0000
1st Qu.	: 23.0000	492.5000
Mean	: 30.9000	668.5000
Median	: 27.5000	610,0000
3rd Qu.	: 38.0000	767.5000
Max	: 52.0000	1390.0000
Total N	: 60.0000	60.0000
Std Dev	: 9.7400	253.6000

Keterangan:

X = Usia Pasien (tahun)

Y = Kadar Monocyte



### Lampiran 3. Makro Program Untuk Menentukan Nilai *Bandwidth* (h) Optimal dengan Fungsi Kernel Gaussian

```
data<-scan("c:/datapatientbcbandwidth.txt")
kernel<-function(t)
{
    if(abs(t)<=1) 1-(abs(t))
    else 0
}
gcv.kernel<-function(data,bandwidth,batas.bawah,batas.atas)
{
    cat("\n *****")
    cat("\n   Kernel Gaussian   ")
    cat("\n *****")
    cat("\n   h           GCV")
    data<-matrix(data,ncol=3)
    n<-nrow(data)
    nonpar<-as.matrix(data[,1])
    X<-as.matrix(data[,2])
    Y<-as.matrix(data[,3])
    h<-batas.bawah
    b<-batas.atas
    iterasi<-((b-h)/bandwidth)+1
    hhasil<-matrix(0,ncol=1,nrow=iterasi)
    gcvhasil<-matrix(0,ncol=1,nrow=iterasi)

    for(k in 1:iterasi)
    {
        z<-matrix(0,ncol=n,nrow=n)
        for(i in 1:n)
        {
            for(j in 1:n)
            {
                z[i,j]<-(-nonpar[i]-nonpar[j])/h
            }
        }
        w<-matrix(0,ncol=n,nrow=n)
        for(i in 1:n)
        {
            for(j in 1:n)
```

### Lampiran 3. (Lanjutan)

```
{
  w[i,j]<-kernel(z[i,j])
}
}
H<-matrix(0,ncol=n,nrow=n)
for(i in 1:n)
{
  for(j in 1:n)
  {
    H[i,j]<-w[i,j]/sum(w[i,])
  }
}
aa<-matrix(0,ncol=n,nrow=n)
x<-matrix(0,ncol=1,nrow=n)
for(i in 1:n)
{
  for(j in 1:n)
  {
    aa[i,j]<-H[i,j]*X[j]
  }
  x[i]<-sum(aa[i,])
  x[i]<-X[i]-x[i]
}
bb<-matrix(0,ncol=n,nrow=n)
y<-matrix(0,ncol=1,nrow=n)
for(i in 1:n)
{
  for(j in 1:n)
  {
    bb[i,j]<-H[i,j]*Y[j]
  }
  y[i]<-sum(bb[i,])
  y[i]<-Y[i]-y[i]
}
x.baru<-cbind(1,x)
Tx.baru<-t(x.baru)
beta<-solve(Tx.baru%*%x.baru,Tx.baru%*%y)
y.baru<-(Y-(cbind(1,X)%*%beta))
f.duga<-H%*%y.baru
atas<-sum((Y-f.duga)^2)/n
bawah<-(-1-(sum(diag(H)))/h)^2
```

### Lampiran 3. (Lanjutan)

```
gcv<-atas/bawah
hhasil[k]<-h
gcvhasil[k]<-gcv

cat("\n ",hhasil[k],"          ",gcvhasil[k], "\n ")
h<-h+bandwidth
}
plot(hhasil,gcvhasil,xlab="Bandwidth",ylab="GCV",main="Plot Bandwidth dan GCV
Kernel Gaussian")
}
```



Lampiran 4. Nilai *Bandwidth* ( $h$ ) dan GCV dengan Fungsi Kernel Gaussian pada Selang Kenaikan  $h=0.05$

```
>gcv.kernel(data, 0.005,0.005,0.2)
```

```
*****
```

```
Kernel Gaussian
```

```
*****
```

$h$	GCV
0.005	1.91373e-06
0.055	0.0002746368
0.105	0.00157305
0.155	0.005650265
0.205	0.01346148
0.255	0.02852584
0.305	0.05135846
0.355	0.09463006
0.405	0.158539
0.455	0.2650818
0.505	0.4141437
0.555	0.6602158
0.605	1.014473
0.655	1.58451
0.705	2.423217
0.755	3.694753
0.805	5.419635
0.855	8.013651
0.905	11.6669
0.955	17.5461
1.005	26.0973
1.055	38.9104
1.105	57.9770
1.155	88.5919
1.205	137.6295
1.255	222.5705
1.305	376.7285
1.355	699.6256

$h$	GCV
2.105	279.2008
2.155	252.0099
2.205	229.9502
2.255	211.7424
2.305	196.4883
2.355	183.5289
2.405	172.4188
2.455	162.8003
2.505	154.4004
2.555	147.0077
2.605	140.4562
2.655	134.6138
2.705	129.3742
2.755	124.6511
2.805	120.3734
2.855	116.4826
2.905	112.9295
2.955	109.6730
3.005	106.6782
3.055	103.9155
3.105	101.3595
3.155	98.9881
3.905	76.8440
3.955	75.9158
4.005	75.0301
4.055	74.1840
4.105	73.3750
4.155	72.6006



Lampiran 4. (Lanjutan)

4.205	71.85879
4.255	71.14749
4.305	70.46491
4.355	69.80936
4.405	69.17929
4.455	68.57326
4.505	67.98992
4.555	67.42804

4.605	66.88647
4.655	66.36414
4.705	65.86005
4.755	65.37327
4.805	64.90293
4.855	64.44822
4.905	64.00838
4.955	63.58271



Lampiran 5. Nilai *Bandwidth* (h) dan GCV dengan Kernel Gaussian pada selang kenaikan  $h=0.005$ .

```
>gcv.kernel(data, 0.005,0.005,0.2)
```

```
*****
```

```
Kernel Gaussian
```

```
*****
```

<i>h</i>	GCV
0.005	1.91373e-06
0.010	7.658249e-06
0.015	1.723856e-05
0.020	3.065966e-05
0.025	4.792657e-05
0.030	6.904432e-05
0.035	9.401791e-05
0.040	0.0001228524
0.045	0.0001555528
0.050	0.0001921242
0.055	0.0002746368
0.060	0.000358459
0.065	0.00044781
0.070	0.0005440805
0.075	0.000647886
0.080	0.0007595442
0.085	0.0008792375
0.090	0.001007081
0.095	0.001143152
0.100	0.001287509

<i>h</i>	GCV
0.105	0.00157305
0.110	0.001873986
0.115	0.002194614
0.120	0.002536508
0.125	0.002900378
0.130	0.003286599
0.135	0.003695397
0.140	0.004126928
0.145	0.004581309
0.150	0.00505864
0.155	0.005650265
0.160	0.006268875
0.165	0.006915029
0.170	0.007589157
0.175	0.0082916
0.180	0.009022637
0.185	0.009782507
0.190	0.01057142
0.195	0.01138957
0.200	0.01223712

Lampiran 6. Nilai *Bandwidth* ( $h$ ) dan GCV dengan Kernel Gaussian pada selang kenaikan  $h=0.001$

```
>gcv.kernel(data, 0.001,0.005,0.05)
```

```
*****
```

```
Kernel Gaussian
```

```
*****
```

$h$	GCV
0.001	1.91373e-06
0.002	2.75601e-06
0.003	3.751563e-06
0.004	4.900427e-06
0.005	6.202642e-06
0.006	7.658249e-06
0.007	9.267287e-05
0.008	1.10298e-05
0.009	1.294582e-05
0.01	1.501539e-05
0.011	1.723856e-05
0.012	1.961535e-05
0.013	2.214582e-05
0.014	2.483e-05
0.015	2.766794e-05
0.016	3.065966e-05
0.017	3.380522e-05
0.018	3.710465e-05
0.019	4.055799e-05
0.02	4.416529e-05
0.021	4.792657e-05
0.022	5.18419e-05
0.023	5.591129e-05

$h$	GCV
0.024	6.01348e-05
0.025	6.451246e-05
0.026	6.904432e-05
0.027	7.373041e-05
0.028	7.857077e-05
0.029	8.356545e-05
0.03	8.871448e-05
0.031	9.401791e-05
0.032	9.947578e-05
0.033	0.0001050881
0.034	0.000110855
0.035	0.0001167764
0.036	0.0001228524
0.037	0.000129083
0.038	0.0001354684
0.039	0.0001420084
0.04	0.0001487032
0.041	0.0001555528
0.042	0.0001625572
0.043	0.0001697166
0.044	0.0001770308
0.045	0.0001845
0.046	0.0001921242

## Lampiran 7. Program Untuk Menentukan Pendugaan $\hat{Y}_i$ dengan Kernel Gaussian

```
data<-scan("c:/datapasiéntbcpenduga.txt")

kernel<-function(t)
{
  if(abs(t)<=1) 1-(abs(t))
  else 0
}
gcv.kernel<-function(data,bandwidth)
{
  cat("\n
*****")
  cat("\n   PendugaanMenggunakan Kernel Gaussian   ")
  cat("\n
*****")
  cat("\n Y   f.dugay.duga(i)   sisaan[i]")
  data<-matrix(data,ncol=3)
  n<-nrow(data)
  nonpar<-as.matrix(data[,1])
  X<-as.matrix(data[,2])
  Y<-as.matrix(data[,3])
  h<-bandwidth

  z<-matrix(0,ncol=n,nrow=n)
  for(i in 1:n)
  {
    for(j in 1:n)
    {
      z[i,j]<--(nonpar[i]-nonpar[j])/h
    }
  }
  w<-matrix(0,ncol=n,nrow=n)
  for(i in 1:n)
  {
    for(j in 1:n)
    {
      w[i,j]<-kernel(z[i,j])
    }
  }
}
```

## Lampiran 7. (Lanjutan)

```
H<-matrix(0,ncol=n,nrow=n)
for(i in 1:n)
{
  for(j in 1:n)
  {
    H[i,j]<-w[i,j]/sum(w[i,])
  }
}
aa<-matrix(0,ncol=n,nrow=n)
x<-matrix(0,ncol=1,nrow=n)
for(i in 1:n)
{
  for(j in 1:n)
  {
    aa[i,j]<-H[i,j]*X[j]
  }
  x[i]<-sum(aa[i,])
  x[i]<-X[i]-x[i]
}
bb<-matrix(0,ncol=n,nrow=n)
y<-matrix(0,ncol=1,nrow=n)
for(i in 1:n)
{
  for(j in 1:n)
  {
    bb[i,j]<-H[i,j]*Y[j]
  }
  y[i]<-sum(bb[i,])
  y[i]<-Y[i]-y[i]
}
x.baru<-cbind(1,x)
Tx.baru<-t(x.baru)
beta<-solve(Tx.baru%*%x.baru,Tx.baru%*%y)
y.baru<--(Y-(cbind(1,X)%*%beta))
f.duga<-H%*%y.baru
y.duga<-(cbind(1,X)%*%beta)+f.duga
sisaan<-Y-y.duga
MSE<-(1/n)*(sum(sisaan^2))
RMSE<-sqrt(MSE)
MAD<-(1/n)*(sum(abs(sisaan)))
```

## Lampiran 7. (Lanjutan)

```
for(i in 1:n)
{
    cat("\n ",usia.pasien[i]," ",Y[i]," ",Y.duga[i]," ",sisaan[i])
}

cat("\n")
cat("\n MSE = ", MSE)
cat("\n RMSE = ", RMSE)
cat("\n MAD = ", MAD)
}
```





Lampiran 8. Hasil Pendugaan  $\hat{Y}_i$  dengan Kernel Gaussian dengan  $Bandwidth (h)=0.005$

```
>gcv.kernel(data,0.005)
```

```
*****
```

PendugaanMenggunakan Kernel Gaussian

```
*****
```

usia. pasien[i]	Y[i]	Y.duga[i]	sisaan[i]	usia. pasien [i]	Y[i]	Y.duga[i]	sisaan[i]
19	240	257,67	-17,67	28	1000	958,72	41,28
20	460	451,31	8,69	28	960	938,34	21,66
21	540	521,14	18,86	29	730	756,15	-26,15
21	580	581	-1	29	490	496,2	-6,2
21	580	577,65	2,35	30	360	340,49	19,51
22	600	602,37	-2,37	30	470	501,49	-31,49
22	640	635,07	4,93	31	490	444,51	45,49
22	650	650	0	32	590	606,75	-16,75
22	700	705,07	-5,07	32	980	986,52	-6,52
22	740	743,52	-3,52	33	1000	994,87	5,13
22	860	858	2	33	740	734,83	5,17
23	1200	1198,01	1,99	34	560	567,59	-7,59
23	1350	1357,9	-7,9	36	400	394,47	5,53
23	980	905,04	74,96	37	380	380	0
23	940	942,19	-2,19	38	380	377,27	2,73
23	760	760,12	-0,12	38	380	364,04	15,96
23	730	732	-2	39	510	461,43	48,57
23	650	647,1	2,9	40	580	517,41	62,59
24	500	503,85	-3,85	41	610	624,72	-14,72
24	490	479,03	10,97	42	620	613,08	6,92
24	370	372,01	-2,01	43	680	689,22	-9,22
24	350	350,19	-0,19	43	770	789	-19
24	380	387	-7	44	780	794,11	-14,11
24	420	415,78	4,22	45	980	963,85	16,15
25	520	522,45	-2,45	50	1390	1285,06	104,94
25	540	550,63	-10,63	50	780	789,06	-9,06
25	570	572,62	-2,62	51	720	675,21	44,79
26	670	686,29	-16,29	51	610	655,05	-45,05
26	1340	1309,38	30,62	52	590	588,72	1,28
27	660	683,9	-23,9	52	570	539,19	30,81

Lampiran 8. (Lanjutan)

MSE = 0.827364

RMSE = 0.873936

MAD = 0.493483

UNIVERSITAS BRAWIJAYA



Lampiran 9. Kadar Monocyte ( $Y_i$ ) Dan Hasil Pendugaan Kadar Monocyte ( $\hat{Y}_i$ )

No	Kadar Monocyte ( $Y_i$ )	( $\hat{Y}_i$ )	No	Kadar Monocyte ( $Y_i$ )	( $\hat{Y}_i$ )
1	240	257,67	31	1000	958,72
2	460	451,31	32	960	858,34
3	540	521,65	33	730	756,15
4	580	577,65	34	490	496,2
5	580	577,65	35	360	340,49
6	600	602,37	36	470	470,49
7	640	645,07	37	490	444,51
8	650	650	38	590	606,75
9	700	705,07	39	980	986,52
10	740	743,52	40	1000	994,87
11	860	861	41	740	734,83
12	1200	1198,01	42	560	567,59
13	1350	1357,9	43	400	394,47
14	980	975,04	44	380	380
15	940	942,19	45	380	377,27
16	760	760,12	46	380	364,04
17	730	732	47	510	531,43
18	650	647,1	48	580	627,41
19	500	503,85	49	610	624,72
20	490	479,03	50	620	613,08
21	370	372,01	51	680	689,22
22	350	350,19	52	770	789
23	380	387	53	780	794,11
24	420	415,78	54	980	963,85
25	520	522,45	55	1390	1085,06
26	540	550,63	56	780	789,06
27	570	572,62	57	720	725,21
28	670	686,29	58	610	665,05
29	1340	1320,38	59	590	588,72
30	660	683,9	60	570	559,19

UNIVERSITAS BRAWIJAYA



UNIVERSITAS BRAWIJAYA

