Pemilihan Model Regresi Terbaik Menggunakan Metode *Akaike's Information Criterion* dan *Schwarz Information Criterion*

M. Fathurahman

Program Studi Ilmu Komputer, FMIPA Universitas Mulawarman
Jl. Barong Tongkok no.5 Kampus Unmul Gn. Kelua Sempaja Samarinda 75119

Abstrak

Analisis regresi seringkali digunakan untuk mengkaji hubungan antara beberapa variabel dan meramal suatu variabel. Agar diperoleh hasil analisis yang optimal, maka diperlukan model regresi terbaik. Beberapa metode dapat digunakan untuk memilih model regresi terbaik, diantaranya adalah dengan metode *Akaike's Information Criterion* (AIC) dan *Schwarz Information Criterion* (SIC). Penelitian ini bertujuan mengkaji pemilihan model regresi terbaik menggunakan metode AIC dan SIC pada kasus faktor-faktor yang mempengaruhi nilai ujian nasional (UNAS) siswa Sekolah Menengah Kejuruan Negeri (SMKN) 1 Samarinda. Berdasarkan metode AIC model regresi terbaik yang dapat digunakan untuk mengetahui hubungan antara rata-rata nilai UNAS siswa SMKN 1 Samarinda dengan rata-rata nilai *tryout* (X_1), nilai kompetensi (X_2) dan rata-rata nilai ujian sekolah (X_3) adalah $Y = -0{,}0094 + 0{,}4541\ X_1 + 0{,}2178\ X_2 + 0{,}3291\ X_3$. Adapun model regresi terbaik menurut metode SIC adalah $Y = 0{,}4749 + -2{,}6174 + + 0{,}5322\ X_1 + \ + 0{,}2636\ X_3$

Kata kunci : Regresi, Model Terbaik, *Akaike's Information Criterion, Schwarz Information Criterion,* UNAS.

Pendahuluan

Analisis regresi merupakan salah satu teknik analisis data dalam statistika yang seringkali digunakan untuk mengkaji hubungan antara beberapa variabel dan meramal suatu variabel (Kutner, Nachtsheim dan Neter, 2004). Jika suatu model regresi digunakan untuk tujuan peramalan, maka diperlukan model terbaik.

Beberapa metode yang dapat digunakan untuk mendapatkan model regresi terbaik, diantaranya adalah dengan metode *Akaike's Information Criterion* (AIC) dan *Schwarz Information Criterion* (SIC) (Widarjono, 2007). Kedua metode tersebut mempunyai kelebihan dibanding menggunakan metode koefisien determinasi (R^2) yang banyak digunakan selama ini. Kelebihan AIC dan SIC adalah terutama pada pemilihan model regresi terbaik untuk tujuan peramalan (*forecasting*), yaitu dapat menjelaskan kecocokan model dengan data yang ada (*in-sample forecasting*) dan nilai yang terjadi di masa mendatang (*out of sample forecasting*). Adapun kelemahan dari metode R^2, diantaranya adalah : (1) metode R^2 hanya digunakan untuk peramalan *in sample* yaitu apakah prediksi model bisa sedekat mungkin dengan data yang ada, (2) tidak ada jaminan bahwa dengan metode R^2 mampu meramalkan nilai di masa mendatang (*out of sample*) dengan baik, (3) metode R^2 harus digunakan dengan syarat variabel tidak bebas (respon) harus sama, (4) nilai R^2 tidak pernah menurun, jika terus ditambahkan variabel prediktor di dalam model walaupun variabel prediktor tersebut kurang atau tidak relevan (Widarjono, 2007).

Oleh karena itu penelitian mengkaji pemilihan model regresi terbaik menggunakan metode AIC dan SIC. Hasil kajian diaplikasikan pada kasus faktor-faktor yang mempengaruhi nilai Ujian Nasional (UNAS) Sekolah Menengah Kejuruan Negeri (SMKN) 1 Samarinda. Penelitian ini dibatasi pada analisis regresi untuk model regresi linier.

Tujuan

Penelitian ini bertujuan mengkaji pemilihan model regresi terbaik menggunakan metode AIC dan SIC pada kasus faktor-faktor yang mempengaruhi nilai UNAS SMKN 1 Samarinda.

Tinjauan Pustaka

Analisis Regresi

Istilah "regresi" pertama kali dikemukakan oleh Sir Francis Galton (1822-1911), seorang antropolog dan ahli meteorologi terkenal dari Inggris. Dalam makalahnya yang berjudul "*Regression towards mediocrity in hereditary stature*", yang dimuat dalam J*ournal of the Anthropological Institute*, volume 15, halaman 246 sampai dengan 263, tahun 1885. Galton menjelaskan bahwa biji keturunan tidak cenderung menyerupai biji induknya dalam hal besarnya, namun lebih medioker (lebih mendekati rata-rata) lebih kecil daripada induknya kalau induknya besar dan lebih besar daripada induknya kalau induknya sangat kecil (Draper dan Smith, 1992).

Dalam analisis regresi, diperlukan suatu model yang digunakan untuk mengetahui hubungan antara variabel tidak bebas (respon) dengan satu atau lebih variabel bebas (prediktor) dan untuk melakukan peramalan terhadap variabel respon. Model regresi dapat diperoleh dengan melakukan estimasi terhadap parameter-parameternya menggunakan metode tertentu. Adapun metode yang dapat digunakan untuk mengestimasi parameter model regresi, khususnya parameter model regresi linier berganda adalah dengan metode kuadrat terkecil (*ordinary least square*) dan metode kemungkinan maksimum (*maximum likelihood*) (Kutner et.al, 2004).

Secara umum model regresi linier berganda dengan k variabel bebas dapat ditulis sebagai berikut (sembiring, 2003):

$$\mathbf{y} = \beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \cdots + \beta_k \mathbf{x}_k + \boldsymbol{\varepsilon} \qquad (1)$$

Bila pengamatan mengenai $\mathbf{y}, \mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_k$ dinyatakan masing-masing dengan $y_i, x_{i1}, x_{i2}, \cdots, x_{ik}$ dan sisa (*error*) ε_i, maka persamaan (1) dapat dituliskan sebagai:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i \qquad (2)$$

dengan $i = 1, 2, \cdots, n$.

Dalam notasi vektor dan matriks, persamaan (2) dapat ditulis menjadi:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \qquad (3)$$

Misalkan lagi

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix}$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \qquad (4)$$

Berdasarkan persamaan (2), (3) dan (4), maka diperoleh model regresi linier berganda sebagai berikut:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (5)$$

dengan:

$\mathbf{y}$: vektor variabel respon berukuran n x 1.
$\mathbf{X}$: matriks variabel prediktor berukuran n x p, untuk $p = k + 1$.
$\boldsymbol{\beta}$: vektor parameter berukuran p x 1.
$\boldsymbol{\varepsilon}$: vektor sisa berukuran n x 1.

Estimasi Parameter

Estimasi parameter ini bertujuan untuk mendapatkan model regresi yang akan digunakan dalam analisis. Beerapa metode dapat digunakan untuk mengestimasi parameter model regresi, salah satu diantaranya adalah metode kuadrat terkecil (*ordinary least square*), yaitu suatu metode estimasi yang meminimumkan jumlah kuadrat sisa (Kutner et.al, 2004).

Menurut Sembiring (2003) estimator yang diperoleh dengan metode kuadrat terkecil mempunyai sifat *best linear unbiased estimator* (BLUE) yaitu estimator yang linier, tidak bias dan mempunyai variansi yang terkecil dari semua estimator linier tidak bias lainnya.

Berdasarkan persamaan (5), maka estimator kuadrat terkecil bagi $\boldsymbol{\beta}$ adalah sebagai berikut:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \qquad (6)$$

Pengujian Parameter

Pengujian parameter berguna untuk mengetahui ada atau tidaknya pengaruh variabel prediktor terhadap variabel respon.

Pengujian parameter dalam analisis regresi linier berganda terdiri dari dua macam, yaitu pengujian parameter secara serentak (simultan) dan secara individu (parsial). Berikut ini dijelaskan kedua jenis pengujian parameter tersebut.

1. Pengujian parameter secara simultan.

   Langkah-langkah pengujian ini adalah sebagai berikut (Sembiring, 2003):

   H_0 : $\beta_1 = \beta_2 = \cdots = \beta_k$
   (Variabel prediktor secara simultan tidak berpengaruh terhadap variabel respon)

   H_1 : Paling tidak ada satu β_j tidak sama dengan nol, dengan j = 1, 2, …, k.
   (Variabel prediktor secara simultan berpengaruh terhadap variabel respon)

   Statistik uji yang digunakan adalah:

   $$F = \frac{\hat{\boldsymbol{\beta}} \mathbf{X}^T \mathbf{Y} / p}{\hat{\boldsymbol{\varepsilon}}^T \hat{\boldsymbol{\varepsilon}} / n - p}$$

   dengan p adalah banyaknya parameter dalam model regresi.

   Daerah kritik:

   H_0 ditolak bila $F \geq F_{(\alpha, p, n-p)}$ atau

   H_0 ditolak bila nilai probabilitas < α.

2. Pengujian parameter secara parsial

   Pengujian ini berguna untuk mengetahui ada atau tidaknya pengaruh masing-masing variabel prediktor terhadap variabel respon. Langkah-langkah pengujian ini adalah sebagai berikut (Sembiring, 2003):

   H_0 : $\beta_j = 0$
   (Variabel prediktor ke-j tidak berpengaruh terhadap variabel respon)

H_1 : $\beta_j \neq 0$

(Variabel prediktor ke-*j* berpengaruh terhadap variabel respon)

dengan $j = 1, 2, \ldots, k$.

Statistik uji:

$$t = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$$

dengan:

$\hat{\beta}_j$: estimator untuk β_j.

$SE(\hat{\beta}_j)$: standar *error* dari β_j.

Daerah kritik :

H_0 ditolak bila $|t| \geq t_{(\frac{\alpha}{2}, n-p)}$ atau H_0 ditolak bila nilai probabilitas < α.

Metode AIC dan SIC

Metode AIC dan SIC adalah metode yang dapat digunakan untuk memilih model regresi terbaik yang ditemukan oleh Akaike dan Schwarz (Grasa, 1989). Kedua metode tersebut didasarkan pada metode *maximum likelihood estimation* (MLE).

Untuk menghitung nilai AIC dan SIC digunakan rumus sebagai berikut:

$$AIC = e^{\frac{2k}{n}} \frac{\sum_{i=1}^{n} \hat{u}_i^2}{n} \tag{7}$$

dengan:

k = jumlah parameter yang diestimasi dalam model regresi

n = jumlah observasi

e = 2,718

u = Sisa (*residual*)

Persamaan (7) dapat juga ditulis sebagai:

$$\ln AIC = \frac{2k}{n} + \ln\left(\frac{\sum_{i=1}^{n} \hat{u}_i^2}{n}\right) \tag{8}$$

$$SIC = n^{\frac{k}{n}} \frac{\sum_{i=1}^{n} \hat{u}_i^2}{n} \tag{9}$$

dengan:

k = jumlah parameter yang diestimasi dalam model regresi

n = jumlah observasi

u = Sisa

Persamaan (9) dapat juga ditulis sebagai:

$$SIC = \left(\frac{k}{n}\right) \ln n + \ln\left(\frac{\sum_{i=1}^{n} \hat{u}_i^2}{n}\right) \tag{10}$$

Menurut metode AIC dan SIC, model regresi terbaik adalah model regresi yang mempunyai nilai AIC dan SIC terkecil (Widarjono, 2007).

Nilai UNAS SMK

Nilai UNAS SMK adalah nilai yang diperoleh siswa setelah melakukan kegiatan pembelajaran selama tiga tahun di jenjang SMK diselenggarakan secara nasional yang meliputi tiga mata uji, yaitu Matematika, Bahasa Inggris dan Bahasa Indonesia. Berikut ini dijelaskan faktor-faktor yang diasumsikan berpengaruh terhadap nilai UNAS SMK.

1. Nilai *Tryout*

   Nilai *Tryout* adalah penilaian yang dilaksanakan secara terpadu dengan kegiatan pemelajaran atau terpisah. Penilaian jenis ini diharapkan mampu meningkatkan standar mengajar, semangat belajar dan akuntabilitas. Hasil penilaian ini dapat digunakan sebagai umpan balik bagi peserta didik untuk mengetahui penguasaan materi sehingga ada memotivasi untuk memperbaiki hasilnya, masukan bagi guru dalam memperbaiki strategi pemelajarannya dan acuan dalam menentukan peserta didik mencapai kompetensi dengan kecepatan belajar yang berbeda-beda meliputi tiga mata pelajaran, yaitu Matematika, Bahasa Inggris dan Bahasa Indonesia.

2. Nilai Uji Kompetensi

   Nilai uji kompetensi adalah suatu penilaian yang dilakukan secara periodik pada satu satuan kompetensi yang pelaksanaannya disesuaikan dengan kegiatan sekolah dan merupakan uji mata diklat terapan dari mata pelajaran Matematika, Bahasa Inggris dan Bahasa Indonesia. Penilaian ini dapat digunakan untuk memantau dan mengendalikan kualitas, pemberian sertifikat, ketercapaian standar bagi SMK.

3. Nilai Ujian Sekolah

   Nilai ujian sekolah adalah nilai yang diperoleh siswa setelah melakukan kegiatan pemelajaran selama tiga tahun di jenjang SMK yang diselenggarakan di tingkat sekolah. Terdiri dari mata pelajaran normatif dan adiktif yang dianggap berpengaruh terhadap nilai UNAS.

Metodologi Penelitian

Sumber data yang digunakan dalam penelitian ini adalah data diperoleh dari tugas akhir (Sahrul, 2009). Kemudian variabel-variabel yang digunakan adalah sebagai berikut:

1. Variabel respon adalah rata-rata nilai UNAS (Y).
2. Variabel prediktor adalah rata-rata nilai tryout (X_1), nilai uji kompetensi (X_2) dan rata-rata nilai ujian sekolah (X_3)

Berdasarkan tujuan penelitian, maka langkah-langkah dalam penelitian ini adalah sebagai berikut:

1. Melakukan estimasi parameter untuk mendapatkan model regresi untuk variabel Y dengan X_1, Y dengan X_2, Y dengan X_3, Y dengan X_1 dan X_2, Y dengan X_1 dan X_3, Y dengan X_2 dan X_3 serta Y dengan X_1, X_2 dan X_3.
2. Melakukan pengujian parameter model regresi yang diperoleh dari 1 secara simultan dan secara parsial.
3. Menghitung nilai AIC dan SIC dari masing-masing model regresi yang diperoleh dari 1.
4. Menentukan nilai AIC dan SIC yang terkecil dari semua model yang diperoleh pada 1.
5. Mendapatkan model regresi terbaik.

Adapun pengolahan data dalam penelitian ini menggunakan bantuan *software* statistika Eviews *under windows* versi 4.

Hasil dan Pembahasan

Estimasi Parameter

Setelah dilakukan estimasi parameter model regresi dengan metode kuadrat terkecil, maka diperoleh hasil seperti pada Tabel 1 berikut.

Tabel 1. Hasil estimasi parameter model regresi

Model	Variabel	Koefisien
1.	Konstanta	2,5779
	X_1	0,6706
2.	Konstanta	0,6921
	X_2	0,9215
3.	Konstanta	-2,9093
	X_3	1,3802
4.	Konstanta	1,6016
	X_1	0,5322
	X_2	0,2636
5.	Konstanta	0,4749
	X_1	0,5499
	X_3	0,3887
6.	Konstanta	-2,6174
	X_2	0,6027
	X_3	0,7371
7.	Konstanta	-0,0094
	X_1	0,4541
	X_2	0,2178
	X_3	0,3291

Berdasarkan Tabel 1 di atas, maka diperoleh model regresi sebagai berikut:

$$Y = 2{,}5779 + 0{,}6706\ X_1 \quad (11)$$
$$Y = 0{,}6921 + 0{,}9215\ X_2 \quad (12)$$
$$Y = -2{,}9093 + 1{,}3802\ X_3 \quad (13)$$
$$Y = 1{,}6016 + 0{,}5322\ X_1 + 0{,}2636\ X_2 \quad (14)$$
$$Y = 0{,}4749 + 0{,}6027\ X_2 + 0{,}7371\ X_3 \quad (15)$$
$$Y = -2{,}6174 + 0{,}5322\ X_1 + 0{,}2636\ X_3 \quad (16)$$
$$Y = -0{,}0094 + 0{,}4541\ X_1 + 0{,}2178\ X_2 + 0{,}3291\ X_3 \quad (17)$$

Pengujian Parameter

Untuk mengetahui ada atau tidaknya pengaruh variabel prediktor terhadap variabel respon, maka dilakukan pengujian parameter secara simultan dan secara parsial. Hasil yang diperoleh adalah seperti pada Tabel 2 dan Tabel 3 berikut.

Tabel 2. Hasil pengujian parameter secara simultan

Model	Variabel	Koefisien	F	Prob.
1.	Konstanta	1,6016	87,84	0,00
	X_1	0,5322		
	X_2	0,2636		
2.	Konstanta	0,4749	89,15	0,00
	X_1	0,5499		
	X_3	0,3887		
3.	Konstanta	-2,6174	50,6	0,00
	X_2	0,6027		
	X_3	0,7371		
4.	Konstanta	-0,0094	63,44	0,00
	X_1	0,4541		
	X_2	0,2178		
	X_3	0,3291		

Berdasarkan Tabel 2 di atas, dapat diketahui bahwa X_1 dan X_2 pada model 1 secara simultan berpengaruh terhadap Y atau dapat dikatakan bahwa rata-rata nilai tryout dan nilai uji kompetensi berpengaruh terhadap Rata-rata nilai UNAS. Hal ini ditunjukkan oleh nilai probabilitasnya (0,00) kurang dari tingkat signifikansi (α = 0,05). Kemudian X_1 dan X_3 pada model 2 secara simultan berpengaruh terhadap Y atau dapat dikatakan bahwa rata-rata nilai tryout dan rata-rata nilai ujian sekolah berpengaruh terhadap Rata-rata nilai UNAS. Hal ini ditunjukkan oleh nilai probabilitasnya (0,00) kurang dari tingkat signifikansi (α = 0,05). Selanjutnya X_2 dan X_3 pada model 3 secara simultan berpengaruh terhadap Y atau dapat dikatakan bahwa nilai kompetensi dan rata-rata nilai ujian sekolah berpengaruh terhadap Rata-rata nilai UNAS. Hal ini ditunjukkan oleh nilai probabilitasnya (0,00) kurang dari tingkat signifikansi (α = 0,05). Kemudian X_1, X_2 dan X_3 pada model 4 secara simultan berpengaruh terhadap Y atau dapat dikatakan bahwa rata-rata nilai *tryout*, nilai kompetensi dan rata-rata nilai ujian sekolah berpengaruh terhadap Rata-rata nilai UNAS. Hal ini ditunjukkan oleh nilai probabilitasnya (0,00) kurang dari tingkat signifikansi (α = 0,05).

Tabel 3. Hasil pengujian parameter secara parsial

Model	Variabel	Koefisien	t	Prob.
1.	Konstanta	2,5779	5,8867	0,0000
	X1	0,6706	12,5843	0,0000
2.	Konstanta	0,6921	0,7669	0,4483
	X2	0,9215	8,1842	0,0000
3.	Konstanta	-2,9093	-1,9087	0,0645
	X3	1,3802	7,2046	0,0000

Berdasarkan Tabel 3 di atas, terlihat bahwa secara parsial X_1 berpengaruh terhadap Y atau dapat dikatakan bahwa rata-rata nilai tryout berpengaruh terhadap rata-rata nilai UNAS. Hal ini ditunjukkan oleh nilai probabilitasnya (0,00) kurang dari tingkat signifikansi (α = 0,05).

Kemudian X_2 pada model 2 secara parsial berpengaruh terhadap Y atau dapat dikatakan bahwa nilai uji kompetensi berpengaruh terhadap rata-rata nilai UNAS. Hal ini ditunjukkan oleh nilai probabilitasnya (0,00) kurang dari tingkat signifikansi (α = 0,05). Selanjutnya X_3 pada model 3 secara parsial berpengaruh terhadap Y atau dapat dikatakan bahwa rata-rata nilai ujian sekolah berpengaruh terhadap rata-rata nilai UNAS. Hal ini ditunjukkan oleh nilai probabilitasnya (0,00) kurang dari tingkat signifikansi (α = 0,05).

Pemilihan Model Regresi Terbaik

Setelah dilakukan estimasi dan pengujian parameter, selanjutnya dilakukan pemilihan model regresi terbaik menggunakan metode AIC dan SIC. Hasil yang diperoleh adalah seperti pada Tabel 4 berikut.

Tabel 4. Hasil pemilihan model regresi terbaik

Model	Variabel	Koefisien	AIC	SIC
1.	Konstanta X1	2,5779 0,6706	0,2125	0,2995
2.	Konstanta X2	0,6921 0,9215	0,8523	0,9393
3.	Konstanta X3	-2,9093 1,3802	1,0122	1,0993
4.	Konstanta X1 X2	1,6016 0,5322 0,2636	0,1565	0,2871
5.	Konstanta X1 X3	0,4749 0,5499 0,3887	0,1441	**0,2747**
6.	Konstanta X2 X3	-2,6174 0,6027 0,7371	0,5953	0,7259
7.	Konstanta X1 X2 X3	-0,0094 0,4541 0,2178 0,3291	0,1177	0,2918

Dari Tabel 4 di atas, terlihat bahwa model regresi terbaik menurut metode AIC adalah model 7 atau model regresi seperti pada persamaan (17). Hal ini ditunjukkan oleh nilai AIC untuk model 7 yang terkecil (0,1177) dibanding nilai AIC untuk model yang lain. Kemudian model regresi terbaik menurut metode SIC adalah model 5 atau model regresi seperti pada persamaan (16). Hal ini ditunjukkan oleh nilai SIC untuk model 5 yang terkecil (0,2747) dibanding nilai SIC untuk model yang lain.

Kesimpulan

Berdasarkan hasil dan pembahasan, maka dapat disimpulkan bahwa model regresi terbaik yang dapat digunakan untuk mengetahui hubungan antara rata-rata nilai UNAS siswa SMKN 1 Samarinda dengan rata-rata nilai *tryout* dan nilai kompetensi dan rata-rata nilai ujian sekolah menurut metode AIC adalah $Y = -0{,}0094 + 0{,}4541\ X_1 + 0{,}2178\ X_2 + 0{,}3291\ X_3$. Kemudian model regresi terbaik menurut metode SIC adalah

$Y = 0{,}4749 + -2{,}6174 + + 0{,}5322\ X_1 + 0{,}2636\ X_3$.

Daftar Pustaka

Grasa, A. A. 1989. *Econometric Model Selection*: *A New Approach*, Kluwer.

Kutner, M.H., Nachtsheim, C.J. dan Neter, J. 2004. *Applied Linear Regression Models.* New York: McGraw-Hill/Irwin.

Sahrul. 2009. Pendekatan Regresi Spline untuk memodelkan dan mengetahui Faktor-faktor yang mempengaruhi Nilai Ujian Nasional Siswa Sekolah Menengah Kejuruan Negeri 1 Samarinda. *Tugas Akhir*. Program Studi Statistika FMIPA Universitas Mulawarman Samarinda.

Sembiring, R.K. 2003. *Analisis Regresi*, Edisi Kedua, Bandung: ITB.

Widarjono, A. 2007. *Ekonometrika*: *Teori dan Aplikasi untuk Ekonomi dan Bisnis*, Yogyakarta: Ekonisia Fakultas Ekonomi Universitas Islam Indonesia.