

## Analisis Faktor Resiko Penderita *Diabetes Mellitus* (DM) Tipe II dengan Menggunakan Regresi *Spline*

(Analysis of Risk Factors for Patients with Type II Diabetes Mellitus (DM) Using Spline Regression)

Linda Fatmawati<sup>1</sup>, Sugiyarto<sup>2\*</sup>

Program Studi Matematika, Fakultas Sains dan Teknologi Terapan, Universitas Ahmad Dahlan Yogyakarta, Jalan Ringroad Selatan, Kragilan, Tamanan, Banguntapan, Bantul, Yogyakarta, 55191, Indonesia

E-mail: [sugiyarto@math.uad.ac.id](mailto:sugiyarto@math.uad.ac.id)

\* Corresponding Author

### ARTICLE INFO

#### Kata Kunci

Regresi Spline  
Diabetes Mellitus

#### Keywords

Spline Regression  
Diabetes Mellitus

### ABSTRACT

*Diabetes Mellitus* (DM) merupakan gangguan metabolisme dari distribusi gula oleh tubuh. Penderita DM tidak bisa memproduksi jumlah insulin dalam jumlah yang cukup, atau tubuh tidak mampu menggunakan insulin secara efektif, sehingga terjadilah kelebihan gula dalam darah. Kelebihan gula yang kronis dalam darah (hiperglikemia) ini menjadi racun bagi tubuh. Klasifikasi atau jenis DM ada bermacam-macam, tetapi di Indonesia yang paling banyak ditemukan adalah DM tipe II. Analisis faktor resiko DM tipe II dengan regresi *spline* mampu mengestimasi data yang tidak memiliki pola tertentu. Hasil permodelan diperoleh kesimpulan bahwa model terbaik dapat dijelaskan melalui model regresi linier kombinasi knot terbobot. Model ini menghasilkan nilai GCV sebesar GCV 24514,2137 dan koefisien determinasi 24,96 %. Dari hasil diatas nampak bahwa variabel yang berpengaruh terhadap gula darah sewaktu adalah hemoglobin, leukosit, tekanan sistol, dan tekanan distol.

Diabetes Mellitus (DM) is a metabolic disorder of sugar distribution by the body. DM sufferers cannot produce sufficient amounts of insulin, or the body is unable to use insulin effectively, resulting in excess sugar in the blood. Chronic excess sugar in the blood (hyperglycemia) is toxic to the body. There are various classifications or types of DM, but in Indonesia the most common type is type II diabetes. Risk factor analysis for type II diabetes mellitus with spline regression is able to estimate data that does not have a specific pattern. The modeling results concluded that the best model can be explained through a linear regression model with the weighted knots combination. This model produces a GCV value of GCV 24514.2137 and a determination coefficient of 24.96%. From the results above, it appears that the variables that affect blood sugar at any time are hemoglobin, leukocytes, systolic pressure, and distole pressure.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



---

## PENDAHULUAN

Penyakit *Diabetes Mellitus* (DM) merupakan penyakit yang mematikan. *International Diabetes Federation* (IDF) menyatakan bahwa di dunia terdapat 366 juta jiwa penderita DM pada tahun 2011 dan diperkirakan menjadi sekitar 522 juta jiwa pada tahun 2030, selain itu pada tahun 2011 sebanyak 4,6 juta jiwa meninggal karena DM.

DM merupakan gangguan metabolisme dari distribusi gula oleh tubuh [1]. Penderita diabetes tidak bisa memproduksi jumlah insulin dalam jumlah yang cukup, atau tubuh tidak mampu menggunakan insulin secara efektif, sehingga terjadilah kelebihan gula dalam darah [2]. Kelebihan gula yang kronis dalam darah (hiperglikemia) ini menjadi racun bagi tubuh.

Klasifikasi atau jenis DM ada bermacam-macam, tetapi di Indonesia yang paling banyak ditemukan adalah DM tipe II. Untuk mengetahui faktor resiko penderita DM tipe II dapat dilakukan dengan analisis regresi *spline*.

Regresi *spline* adalah suatu pendekatan ke arah pengepasan data dengan tetap memperhitungkan kemulusan kurva. Regresi *spline* memungkinkan untuk berbagai macam orde sehingga dapat dibentuk linier, kuadrat, kubik maupun orde  $m$  [3]. Bentuk estimator *spline* sangat dipengaruhi oleh parameter penghalus, lokasi titik knot, dan banyaknya titik knot. Pemilihan parameter penghalus optimal dalam regresi *spline* pada hakekatnya merupakan pemilihan lokasi titik knot [4]. Knot merupakan titik perpaduan bersama dimana terdapat perubahan perilaku pola pada interval berlainan.

Oleh karena itu peneliti tertarik untuk menerapkan regresi *spline* untuk analisis faktor resiko penyakit DM tipe II. Tujuan utamanya adalah untuk mengetahui secara nyata penyebab dari DM tipe II tersebut sehingga dapat dilakukan pencegahan dini agar dapat mengurangi resiko menderita penyakit DM tipe II.

### Diabetes Mellitus

Penyakit *Diabetes Mellitus* adalah penyakit dengan gejala kadar gula darah yang tinggi yang disebabkan tubuh tidak lagi memiliki hormon insulin atau insulin tidak dapat bekerja sebagaimana mestinya.

*Diabetes Mellitus* (DM) tipe II yaitu diabetes yang tidak tergantung pada insulin. Biasanya terjadi sekitar 90%-95% dari penderita diabetes secara keseluruhan. DM tipe II terjadi akibat penurunan sensitivitas terhadap insulin atau akibat penurunan jumlah produksi insulin [5].

Perkumpulan Endokrinologi Indonesia (PERKENI) 2006 menetapkan batasan-batasan diagnosa yang menyatakan orang tersebut menderita penyakit *Diabetes Mellitus* bila pemeriksaan kadar gula darah puasa  $>126$  mg/dl atau kadar gula darah sewaktu  $>200$  mg/dl. Kadar gula darah untuk orang normal sebelum makan pagi berkisar antara 70 – 110 mg/dl dan 2 jam sesudah makan pagi berkisar 120 – 140 mg/dl [6].

Berikut hubungan antara kadar gula darah tipe II dengan kadar hemoglobin, leukosit, dan tekanan Darah [5]

1. Anemia didefinisikan sebagai penurunan dalam kapasitas transportasi oksigen dalam darah. Hal ini dapat timbul jika ada terlalu sedikit hemoglobin yang beredar atau hemoglobin yang berfungsi [7]. Anemia telah dikaitkan dengan terjadinya komplikasi mikrovaskuler dan makrovaskuler pada diabetes. Salah satu komplikasi DM tipe II adalah penyakit ginjal kronis yang mana insidennya juga semakin tinggi di dunia.
2. Leukositosis adalah keadaan dengan jumlah sel darah putih (leukosit) dalam darah meningkat, melebihi nilai normal. Jumlah leukosit per mikroliter darah, pada orang dewasa normal adalah 4000-11000 sel [8]. Penderita DM ditemukan lebih banyak kuman dan jamur di tubuhnya. Pada keadaan normal kuman-kuman masuk ke dalam tubuh akan di lawan dan dibunuh oleh pasukan pertahanan tubuh yaitu leukosit. Pada waktu gula darah tinggi atau lebih dari 200 mg/dl, kekuatan sel-sel darah putih untuk pergerakan, penempelan fagositosis sel dan kemampuan membunuh kuman akan berkurang sehingga kuman yang masuk lebih sukar untuk dibunuh dan justru terus berkembang biak menjadi infeksi. Salah satu faktor yang mungkin menyebabkan DM tipe II adalah infeksi virus, seperti gondongan, rubella, atau sitomegalovirus [2].
3. Dikatakan tekanan darah tinggi (hipertensi) jika pada saat duduk tekanan sistolik mencapai 140 mmHg atau lebih, atau tekanan diastolik mencapai 90 mmHg atau lebih, atau keduanya. Apabila kondisi hipertensi seseorang dibiarkan tanpa perawatan maka kondisi ini dapat menyebabkan

penebalan pembuluh darah arteri yang menyebabkan diameter pembuluh darah menjadi menyempit. Hal ini akan menyebabkan proses pengangkutan glukosa dari dalam darah menjadi terganggu [9].

Regresi Spline

Regresi spline merupakan suatu pendekatan ke arah pengepasan data dengan tetap memperhitungkan kemulusan kurva[10]. Spline merupakan model polinomial yang tersegmen. Sifat tersegmen inilah yang memberikan fleksibilitas yang lebih baik daripada model polinomial biasa. Sifat ini memungkinkan model regresi spline menyesuaikan diri secara efektif terhadap karakteristik lokal dari data [11].

Bentuk estimator spline sangat dipengaruhi oleh parameter penghalus  $\lambda$ , lokasi titik knot, dan banyaknya titik knot. Pemilihan parameter penghalus  $\lambda$  optimal dalam regresi spline pada hakekatnya merupakan pemilihan lokasi titik knot [4]. Knot merupakan titik perpaduan bersama dimana terdapat perubahan perilaku pola pada interval berlainan .

Fungsi regresi yang didekati oleh fungsi spline dan melibatkan satu variabel respon dan banyak variabel prediktor disebut sebagai regresi spline multivariabel. Kurva regresi  $f$  yang dihipotesiskan oleh fungsi spline multivariabel dituliskan sebagai:

$$f(x_{1i}, x_{2i}, \dots, x_{pi}) = \sum_{j=1}^p f(x_{ji}) \tag{3.1}$$

dengan

$$f(x_{ji}) = \alpha_{00} + \sum_{h=1}^q \alpha_{hj} x_{ji}^h + \sum_{l=1}^m \beta_{lj} (x_{ji} - \lambda_{lj})_+^q$$

$$= \alpha_{00} + \alpha_{1j} x_{ji} + \dots + \alpha_{qj} x_{ji}^q + \beta_{1j} (x_{ji} - \lambda_{1j})_+^q + \dots + \beta_{mj} (x_{ji} - \lambda_{mj})_+^q, \tag{3.2}$$

dimana  $j = 0, 1, 2, \dots, p$ . Model regresi fungsi tersebut berbentuk

$$y_i = \alpha_{00} + \alpha_{11} x_{1i} + \dots + \alpha_{q1} x_{1i}^q + \beta_{11} (x_{1i} - \lambda_{11})_+^q + \dots + \beta_{m1} (x_{1i} - \lambda_{m1})_+^q + \dots + \alpha_{1j} x_{ji} + \dots + \alpha_{qj} x_{ji}^q + \beta_{1j} (x_{ji} - \lambda_{1j})_+^q + \dots + \beta_{mj} (x_{ji} - \lambda_{mj})_+^q + \varepsilon_i \tag{3.3}$$

dengan

$$(x_{ji} - \lambda_{lj})_+^q = \begin{cases} (x_{ji} - \lambda_{lj})^q, & x_{ji} \geq \lambda_{lj} \\ 0, & x_{ji} < \lambda_{lj} \end{cases}$$

Atau  $y = X(\lambda_{11}, \dots, \lambda_{m1} : \dots : \lambda_{1j}, \dots, \lambda_{mj})\theta + \varepsilon$ , dengan  $y = [y_1, \dots, y_n]'$ ,

$$X(\lambda_{11}, \dots, \lambda_{m1} : \dots : \lambda_{1j}, \dots, \lambda_{mj}) =$$

$$\begin{bmatrix} x_{11} & \dots & x_{11}^q (x_{11} - \lambda_{11})_+^q & \dots & (x_{11} - \lambda_{m1})_+^q & : & \dots & : & x_{j1} & \dots & x_{j1}^q (x_{j1} - \lambda_{1j})_+^q & \dots & (x_{j1} - \lambda_{mj})_+^q \\ x_{12} & \dots & x_{12}^q (x_{12} - \lambda_{11})_+^q & \dots & (x_{12} - \lambda_{m1})_+^q & : & \dots & : & x_{j2} & \dots & x_{j2}^q (x_{j2} - \lambda_{1j})_+^q & \dots & (x_{j2} - \lambda_{mj})_+^q \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ x_{1n} & \dots & x_{1n}^q (x_{1n} - \lambda_{11})_+^q & \dots & (x_{1n} - \lambda_{m1})_+^q & : & \dots & : & x_{jn} & \dots & x_{jn}^q (x_{jn} - \lambda_{1j})_+^q & \dots & (x_{jn} - \lambda_{mj})_+^q \end{bmatrix}$$

$\theta = (\theta'_1, \dots, \theta'_j)'$  dengan  $\theta'_1 = (\alpha_{11}, \dots, \alpha_{q1}, \beta_{11}, \dots, \beta_{m1})$  dan  $\theta'_j = (\alpha_{1j}, \dots, \alpha_{qj}, \beta_{1j}, \dots, \beta_{mj})$  dan  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)'$

Penaksiran untuk  $\theta = (\theta'_1, \dots, \theta'_j)'$  diperoleh dengan menggunakan metode least square. Penaksir untuk  $\theta$ , yaitu  $\hat{\theta}$  diberikan oleh

$$\hat{\theta} = \left( X'(\lambda_{11}, \dots, \lambda_{m1}, \dots, \lambda_{1j}, \dots, \lambda_{mj}) X(\lambda_{11}, \dots, \lambda_{m1}, \dots, \lambda_{1j}, \dots, \lambda_{mj}) \right)^{-1} X'(\lambda_{11}, \dots, \lambda_{m1}, \dots, \lambda_{1j}, \dots, \lambda_{mj}) y \tag{3.4}$$

dimana  $\hat{\theta}'_1 = (a_{11}, \dots, a_{q1}, b_{11}, \dots, b_{m1})$  dan  $\hat{\theta}'_j = (a_{1j}, \dots, a_{qj}, b_{1j}, \dots, b_{mj})$

Penaksir kurva  $f(x)$  diperoleh dari:

$$\hat{f}(x) = H(\lambda_{11}, \dots, \lambda_{m1}, \dots, \lambda_{1j}, \dots, \lambda_{mj})y$$

Matrik  $H(\lambda_{11}, \dots, \lambda_{m1}, \dots, \lambda_{1j}, \dots, \lambda_{mj})$  pada persamaan merupakan matrik yang diberikan oleh persamaan:

$$H(\lambda_{11}, \dots, \lambda_{m1}, \dots, \lambda_{1j}, \dots, \lambda_{mj}) = \left( X'(\lambda_{11}, \dots, \lambda_{m1}, \dots, \lambda_{1j}, \dots, \lambda_{mj}) X(\lambda_{11}, \dots, \lambda_{m1}, \dots, \lambda_{1j}, \dots, \lambda_{mj}) \right)^{-1} X'(\lambda_{11}, \dots, \lambda_{m1}, \dots, \lambda_{1j}, \dots, \lambda_{mj})$$

[10]

#### Pemilihan Titik Knot Optimal

Bentuk estimator spline sangat dipengaruhi oleh lokasi dan banyaknya titik knot( $\lambda$ ). Parameter  $\lambda$  merupakan pengontrol dalam melihat kemulusan kurva dan keseimbangan kurva terhadap data sehingga parameter  $\lambda$  yang optimal menjadi penting berkaitan dengan diperolehnya estimator yang optimal. Pemilihan  $\lambda$  optimal memenuhi beberapa kriteria tertentu misalnya nilai *Mean Square Error* (MSE) atau *Generalized Least Square* (GCV) yang minimum.

$$GCV(\lambda) = \frac{MSE(\lambda)}{(n^{-1} \text{trace}[1-H(\lambda)])^2} \quad (3.5)$$

dengan

$$MSE(\lambda) = \frac{SSE}{n} = \frac{1}{n} \sum_{j=1}^n (y_j - \hat{f}_{\lambda_j})^2 \quad (3.6)$$

N : jumlah pengamatan

I : matriks identitas

H : matriks  $X_{\lambda}(X_{\lambda}^T X_{\lambda})^{-1} X_{\lambda}^T$

[4].

## METODE

#### Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diambil dari RS PKU Muhammadiyah Bantul. Populasi penelitian adalah seluruh data rekam medis pasien DM tipe II. Sampel yang diambil yaitu rekam medis pasien DM tipe II yang tercatat pada periode Januari-Desember 2013. Berdasarkan data tersebut, diketahui bahwa terdapat 254 orang yang melakukan tes kadar gula darah, tes hemoglobin, tes leukosit dan tes tekanan darah.

#### Variabel Penelitian

Terdapat empat variabel prediktor dan satu variabel respon untuk analisis regresi *spline* dalam penelitian ini, yaitu:

- Kadar gula darah sewaktu ( $y$ )
- Kadar hemoglobin ( $x_1$ )
- Kadar leukosit ( $x_2$ )
- Tekanan darah sistol ( $x_3$ )
- Tekanan darah distol ( $x_4$ )

#### Langkah Penelitian

Langkah-langkah yang dilakukan adalah sebagai berikut:

1. Membuat *scatter plot* antara variabel prediktor dan variabel respon untuk mengetahui perilaku pola data.
2. Memodelkan kadar gula darah sewaktu dengan regresi *spline* linier 1 knot, 2 knot, 3 knot, dan kombinasi ketiga titik knot.
3. Memilih titik optimal menggunakan metode *Generalized Cross Validation* (GCV) dimana GCV optimal berkaitan dengan GCV minimum.
4. Memodelkan kadar gula darah sewaktu menggunakan regresi *spline* dengan knot optimal.

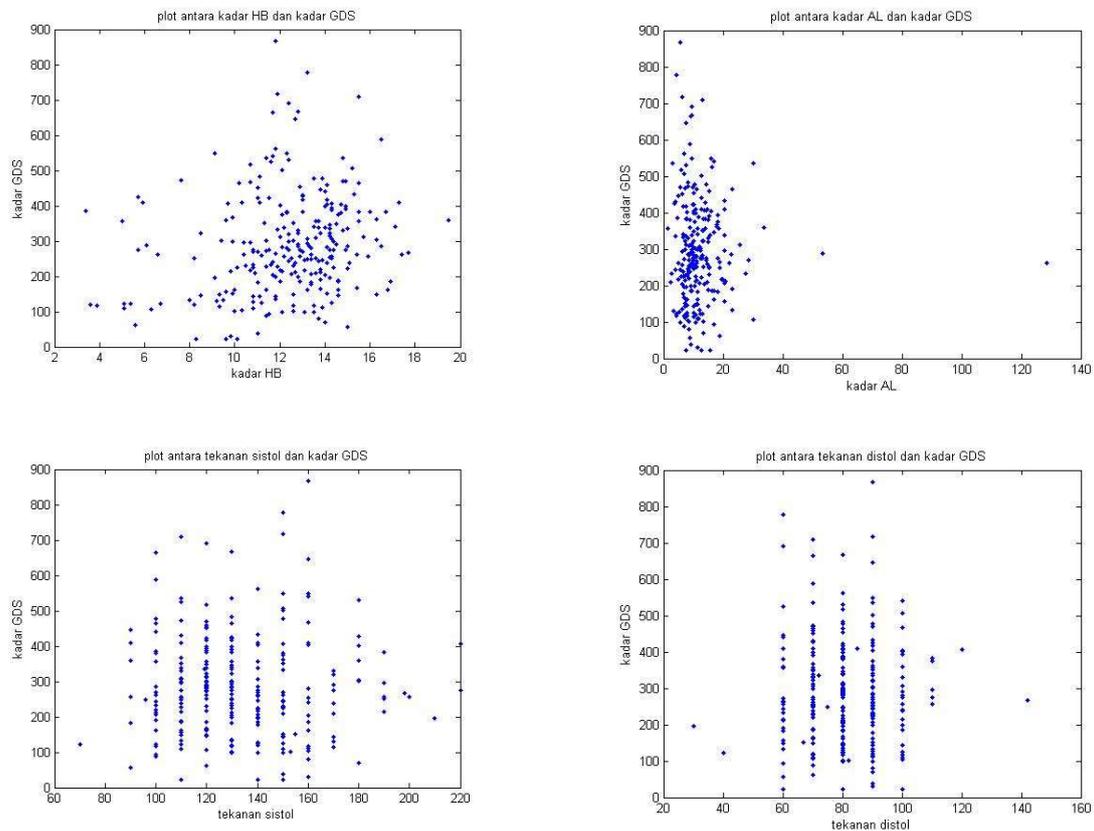
5. Melakukan uji signifikansi parameter dengan menggunakan uji serentak dan uji parsial.
6. Melakukan uji asumsi residual, yaitu uji asumsi identik menggunakan uji glejser, uji independen menggunakan uji Durbin-Watson, dan uji distribusi normal menggunakan uji Kolmogorof-Smirnov. Apabila terdapat heteroskedastisitas, maka digunakan regresi *spline* terboboti.
7. Melakukan interpretasi model.
8. Menarik kesimpulan.

### HASIL DAN PEMBAHASAN

Scatterplot Kadar Gula Darah Sewaktu dengan Faktor yang Mempengaruhinya

Bentuk pola hubungan fungsional antara variabel prediktor dengan variabel respon dapat diperkirakan dengan membuat diagram pencar (scatterplot) yang memuat informasi tentang kedua hubungan tersebut. Bentuk pola hubungan antara variabel respon dengan variabel prediktor disajikan pada gambar 1

Gambar 1 menunjukkan bahwa pola hubungan antara variabel respon dengan variabel prediktor memiliki pola yang menyebar, sehingga tidak memiliki kecenderungan membentuk suatu pola tertentu atau tidak mengikuti pola tertentu. Oleh karena itu sulit digunakan pemodelan dengan pendekatan regresi parametrik. Selanjutnya pola data akan didekati dengan menggunakan regresi nonparametrik *spline* linier.



Gambar 1. hubungan variabel prediktor dan variabel respon

#### Regresi *Spline* Linier

Permodelan dengan menggunakan regresi *spline* dilakukan dengan memodelkan 1, 2, 3, dan kombinasi ketiga titik knot. Dari hasil permodelan dihitung pula nilai GCV dari masing-masing model regresi *spline* linier.

Tabel 1. Perbandingan Nilai GCV dari Keempat Model *Spline*

Model	GCV Minimum
1 knot	20481,4224
2 knot	20871,8399
3 knot	20905,4843
Kombinasi Knot	20184,3877

Berdasarkan perbandingan GCV yang ada pada tabel 1 maka dapat diperoleh model terbaik yaitu dengan menggunakan *spline* linier dengan kombinasi knot yaitu nilai GCV sebesar 20184,3877. Nilai GCV ini paling kecil bila dibandingkan dengan nilai GCV dari *spline* linier dengan 1 knot, 2 knot dan 3 knot yang masing-masing GCV-nya 20481,4224, 20871,8399 dan 20905,4843 dimana regresi *spline* linier letak titik knot berada di  $\lambda_{11} = 10,8$ ,  $\lambda_{21} = 11,7$ ,  $\lambda_{31} = 13,5$ ,  $\lambda_{12} = 3,8$ ,  $\lambda_{22} = 6,7$ ,  $\lambda_{32} = 9,7$ ,  $\lambda_{13} = 120$ , dan  $\lambda_{14} = 70$  dengan koefisien determinasi 11,26%. Hasil dari estimasi parameter dengan menggunakan kombinasi knot adalah sebagai berikut:

$$\hat{y}_i = -43,3547 + 8,2571x_{1i} + 128,2617(x_{1i} - 10,8)_+ - 180,9391(x_{1i} - 11,7)_+ + 65,0924(x_{1i} - 13,5)_+ + 25,5510x_{2i} - 72,0196(x_{2i} - 3,8)_+ + 51,8546(x_{2i} - 6,7)_+ - 4,5508(x_{2i} - 9,7)_+ + 0,8992x_{3i} - 0,9665(x_{3i} - 120)_+ + 1,6264x_{4i} - 2,3420(x_{4i} - 70)_+ + e_i$$

Uji Signifikansi Parameter

Terdapat dua uji estimasi parameter yang dilakukan yaitu uji parameter secara serentak dan uji secara parsial. Uji serentak merupakan pengujian parameter model regresi secara bersamaan. Sedangkan uji parsial adalah pengujian parameter model regresi secara satu persatu.

Dalam uji serentak dilakukan uji hipotesis sebagai berikut:

$$H_0: a_{11} = a_{12} = a_{13} = a_{14} = b_{11} = b_{12} = b_{13} = b_{14} = 0$$

$$H_1: \text{Minimal ada satu } a_{ji} \neq 0 \text{ atau } b_{ji} \neq 0$$

Tabel 2. *Anova*

Sumber Variasi	df	SS	MS	$F_{hitung}$	$F_{tabel}$
Regresi	12	585400,5464	48783,3789		
Error	241	4615470,1640	19151,3285	2,5473	1,7925
Total	253	5200870,7068			

Pada Tabel 2 dapat dilihat bahwa nilai  $F_{hitung}$  sebesar 2,5473. Pada pengujian ini taraf signifikansinya adalah 5% dan  $F_{0,05(12,241)} = 1,7925$ . Daerah penolakan  $H_0$  dilakukan apabila  $F_{hitung} > F_{tabel}$ . Sehingga dapat ditarik kesimpulan bahwa di dalam parameter pada model ini terdapat minimal satu parameter yang signifikan.

Untuk melihat parameter mana yang signifikan atau tidak dalam model, dilakukan uji parsial. Rumusan hipotesisnya sebagai berikut:

$$H_0: a_{11} = 0, \quad H_1: a_{11} \neq 0$$

$$\vdots$$

$$H_0: a_{14} = 0, \quad H_1: a_{14} \neq 0$$

$$H_0: b_{11} = 0, \quad H_1: b_{11} \neq 0$$

$$\vdots$$

$$H_0: b_{14} = 0, \quad H_1: b_{14} \neq 0$$

$H_0$  ditolak apabila  $|t_{hitung}| > t_{\alpha/2(n-m-1)}$ . Dengan nilai  $n$  adalah jumlah observasi dan nilai  $m$  adalah jumlah parameter. Dengan taraf signifikansi = 5% , maka diperoleh  $t_{0,025(241)}=1,970$

Tabel 3. Uji Parsial

Parameter	Estimasi	$t_{hitung}$	Keputusan
$a_{11}$	8,2571	1,5482	Tidak Signifikan
$b_{11}$	128,2617	11,1807	Signifikan
$b_{21}$	-180,9391	-13,4343	Signifikan
$b_{31}$	65,0924	3,0046	Signifikan
$a_{12}$	25,5510	4,5271	Signifikan
$b_{12}$	-72,0196	-10,5077	Signifikan
$b_{22}$	51,8546	6,0858	Signifikan
$b_{32}$	-4,5508	-0,3933	Tidak Signifikan
$a_{13}$	0,8992	0,5219	Tidak Signifikan
$b_{13}$	-0,9665	-0,2027	Tidak Signifikan
$a_{14}$	1,6264	0,7413	Tidak Signifikan
$b_{14}$	-2,3420	-0,4282	Tidak Signifikan

Tabel 3 menunjukkan terdapat 6 parameter yang menghasilkan  $|t_{hitung}| > t_{\alpha/2(n-m-1)}$ , yaitu parameter  $b_{11}$ ,  $b_{21}$  dan  $b_{31}$  dari variabel anemia dan parameter  $a_{12}$ ,  $b_{12}$ ,  $b_{22}$  dari variabel leukositosis. Sehingga dapat dikatakan bahwa kedua variabel tersebut memberikan pengaruh yang signifikan terhadap model. Sehingga model regresi *spline* dari faktor-faktor yang mempengaruhi resiko DM tipe II adalah sebagai berikut

$$\hat{y}_i = 128,2617(x_{1i} - \lambda_{11})_+ - 180,9391(x_{1i} - \lambda_{21})_+ + 65,0924(x_{1i} - \lambda_{31})_+ + 25,5510x_{2i} - 72,0196(x_{2i} - \lambda_{12})_+ + 51,8546(x_{2i} - \lambda_{22})_+$$

Uji Asumsi Residual

Uji asumsi residual pada regresi *spline* serupa dengan uji asumsi residual pada regresi parametrik yaitu residual harus identik, independen dan berdistribusi normal. Pengujian identik bertujuan untuk mengetahui apakah varians dari residual homogen. Uji hipotesis statistik yang digunakan untuk uji identik adalah uji *Glejser*. Pengujian ini dilakukan dengan cara meregresikan harga mutlak residual dengan variabel prediktor.

Tabel 4. Uji *Glejser*

Sumber Variasi	df	SS	MS	$F_{hitung}$	$F_{tabel}$
Regresi	12	168530,6197	14044,2183		
Error	241	1606480,1833	6665,8929	2,1069	1,7925
Total	253	1775010,8030			

Dengan menggunakan taraf signifikansi  $\alpha = 0,05$  didapatkan nilai  $F_{0,05;12;241} = 1,7925$ . Pada tabel 4.17 dapat dilihat bahwa nilai  $F_{hitung}$  sebesar  $2,1069 > F_{tabel}$ . sehingga keputusannya adalah tolak  $H_0$  yang artinya varians residual tidak homogen (terjadi heteroskedastisitas).

Selanjutnya dilakukan pengujian asumsi independen bertujuan untuk mengetahui ada tidaknya korelasi antar residual. Salah satu cara untuk mendeteksi ada tidaknya korelasi antar residual adalah dengan uji Durbin-Watson.

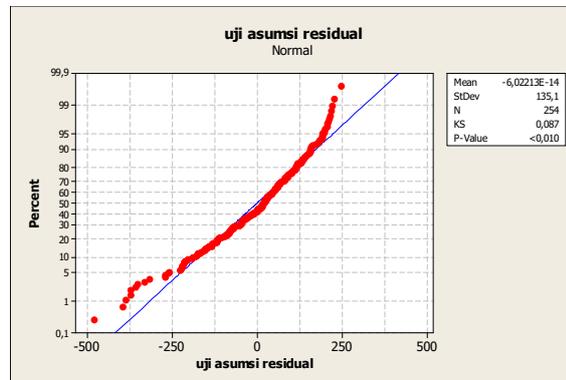
Tabel 5. Uji Durbin-Watson

$d$	$d_L$	$d_U$	$4 - d_U$	$4 - d_L$
2,0010	1,69312	1,89393	2,10607	2,30688

Pada Tabel 5 nilai  $d$  di antara  $d_U$  dan  $4 - d_U$ . Sehingga tidak terjadi autokorelasi.

Metode uji normalitas menggunakan uji *Kolmogorov-Smirnov*. Pengujian distribusi normal pada model regresi digunakan untuk menguji apakah nilai residual yang dihasilkan berdistribusi secara

normal atau tidak. Model regresi yang bagus memiliki nilai residual yang berdistribusi normal. Dari Gambar 2 diperoleh informasi bahwa  $p - value < 0,010$ . Sehingga keputusan yang diambil adalah tolak  $H_0$  yang berarti residual tidak berdistribusi normal.



Gambar 2. Uji Distribusi Normal

Dari pengujian asumsi residual dapat diketahui bahwa terjadi heteroskedastisitas. Untuk mengatasi masalah ini, data perlu didekati dengan regresi *spline* terbobot.

Regresi *Spline* Linier Terbobot

Kasus heteroskedastisitas menyebabkan lebih besarnya variansi dari taksiran. Hal ini tentunya akan berpengaruh pada uji hipotesis yang dilakukan (uji  $F$  dan uji  $t$ ) karena uji tersebut menggunakan besarnya variansi taksiran, akibatnya kedua uji tersebut menjadi kurang akurat. Estimasi parameter dengan menggunakan *Weighted Least Square* (WLS) dapat mengatasi kasus heteroskedastisitas dan tetap diperoleh estimasi parameter yang tidak bias namun tidak efisien.

Penentuan nilai pembobot menjadi permasalahan yang penting karena pembobot yang digunakan idealnya menghasilkan residual yang identik dan memenuhi asumsi IIDN (Identik, Independen, dan Distribusi Normal). Penentuan bobot dilakukan dengan coba-coba (trial error) menggunakan fungsi prediktornya. Dengan menggunakan kombinasi knot 3311, yaitu 10,8 11,7 13,5 3,8 6,7 9,7 120 dan 70. Model *spline* ini ditulis menjadi

$$\hat{y}_i = a_{00} + a_{11}x_{1i} + b_{11}(x_{1i} - 10,8)_+ + b_{21}(x_{1i} - 11,7)_+ + b_{31}(x_{1i} - 13,4)_+ + a_{12}x_{2i} + b_{12}(x_{2i} - 3,8)_+ + b_{22}(x_{2i} - 6,7)_+ + b_{32}(x_{2i} - 9,5)_+ + a_{13}x_{3i} + b_{13}(x_{3i} - 120)_+ + a_{14}x_{4i} + b_{14}(x_{4i} - 70)_+ + e_i$$

Dengan menggunakan bobot  $\frac{1}{x_1^4}$ , estimasi parameter yang terbentuk disajikan dalam Tabel 6 diperoleh nilai GCV 24514,2137 dengan koefisien determinasi 24,96 %.

Tabel 6. Titik – Titik Knot dan Estimasi Parameter Regresi *Spline* Linier

$x_j$	Knot	Estimasi parameter
1	10,8	$a_{11} = 13,8465$
	11,7	$b_{11} = 94,2316$
	13,4	$b_{21} = -150,1617$
		$b_{31} = 58,7381$
2	3,8	$a_{12} = -41,8248$
	6,7	$b_{12} = 30,2921$
	9,5	$b_{22} = 0,0620$
		$b_{32} = 12,3023$
3		$a_{13} = -7,6663$
	180	$b_{13} = 7,2464$
4		$a_{14} = 7,2979$
	70	$b_{14} = -3,7579$
Perpotongan		$a_{00} = 674,2593$

Uji Signifikansi Parameter Regresi *Spline* Terbobot

Dalam uji serentak dilakukan uji hipotesis sebagai berikut:

$$H_0: a_{11} = a_{12} = a_{13} = a_{14} = b_{11} = b_{12} = b_{13} = b_{14} = 0$$

$$H_1: \text{Minimal ada satu } a_{ji} \neq 0 \text{ atau } b_{ji} \neq 0$$

Tabel 7. Uji Serentak

Sumber Variasi	df	SS	MS	$F_{hitung}$	$F_{tabel}$
Regresi	12	1298949,1660	108245,7638		
Error	242	5605551,3522	23259,5492	4,6538	1,7925
Total	254	5200870,7068			

Pada tabel 7 dapat dilihat bahwa nilai  $F_{hitung}$  sebesar 4,6538. Pada pengujian ini taraf signifikansinya adalah 5% dan  $F_{0,05(12,241)} = 1,7925$ . Daerah penolakan  $H_0$  dilakukan apabila  $F_{hitung} > F_{tabel}$ . Sehingga dapat ditarik kesimpulan bahwa di dalam parameter pada model ini terdapat minimal satu parameter yang signifikan.

Untuk melihat parameter mana yang signifikan atau tidak dalam model, dilakukan uji parsial. Rumusan hipotesisnya sebagai berikut:

$$H_0: a_{11} = 0, \quad H_1: a_{11} \neq 0$$

$$\vdots$$

$$H_0: a_{14} = 0, \quad H_1: a_{14} \neq 0$$

$$H_0: b_{11} = 0, \quad H_1: b_{11} \neq 0$$

$$\vdots$$

$$H_0: b_{14} = 0, \quad H_1: b_{14} \neq 0$$

$H_0$  ditolak apabila  $|t_{hitung}| > t_{\alpha/2(n-m-1)}$ . Dengan nilai  $n$  adalah jumlah observasi dan nilai  $m$  adalah jumlah parameter. Dengan taraf signifikansi = 5%, maka diperoleh  $t_{0,025(241)} = 1,970$

Tabel 8. Uji Parsial

Parameter	Estimasi	$t_{hitung}$	Keputusan
$a_{11}$	13,8465	2,3558	Signifikan
$b_{11}$	94,2316	7,4536	Signifikan
$b_{21}$	-150,1617	-10,1167	Signifikan
$b_{31}$	58,7381	2,4602	Signifikan
$a_{12}$	-41,8248	-6,7243	Signifikan
$b_{12}$	30,2921	4,0103	Signifikan
$b_{22}$	0,0620	0,0066	Tidak Signifikan
$b_{32}$	12,3023	0,9648	Tidak Signifikan
$a_{13}$	-7,6663	-4,0372	Signifikan
$b_{13}$	7,2464	1,3788	Tidak Signifikan
$a_{14}$	7,2979	3,0185	Signifikan
$b_{14}$	-3,7579	-0,6235	Tidak Signifikan

Tabel 8 menunjukkan terdapat 8 parameter yang menghasilkan  $|t_{hitung}| > t_{\alpha/2(n-m-1)}$ , yaitu parameter  $a_{11}$ ,  $b_{11}$ ,  $b_{21}$  dan  $b_{31}$  dari variabel hemoglobin, parameter  $a_{12}$  dan  $b_{12}$  dari variabel leukosit, parameter  $a_{13}$  dari tekanan sistol, dan parameter  $a_{14}$  dari tekanan distol. Sehingga dapat dikatakan bahwa keempat variabel tersebut memberikan pengaruh yang signifikan terhadap model. Model regresi *spline* dari faktor-faktor yang mempengaruhi resiko DM tipe II adalah sebagai berikut

$$\hat{y}_i = 13,8465x_{1i} + 94,2316(x_{1i} - 10,8)_+ - 150,1617(x_{1i} - 11,7)_+ + 58,7381(x_{1i} - 13,4)_+ - 41,8248x_{2i} + 30,2921(x_{2i} - 3,8)_+$$

$$-7,6663x_{3i} + 7,2979x_{4i}$$

Uji Asumsi Residual Regresi *Spline* Terbobot

Uji hipotesis statistik yang digunakan untuk uji identik adalah uji *Glejser*.

Sumber Variasi	df	SS	MS	$F_{hitung}$	$F_{tabel}$
Regresi	12	196797,8023	16399,8168		
Error	241	2214926,1964	9190,5651	1.7844	1,7925
Total	253	2319804,8746			

Dengan menggunakan taraf signifikansi  $\alpha = 0,05$  didapatkan nilai  $F_{0,05;12;242} = 1,7925$ . Pada Tabel 9 dapat dilihat bahwa nilai  $F_{hitung}$  sebesar  $1.7844 < F_{tabel}$ . Sehingga keputusannya adalah gagal tolak  $H_0$  yang artinya varians residual homogen (tidak terjadi heteroskedastisitas).

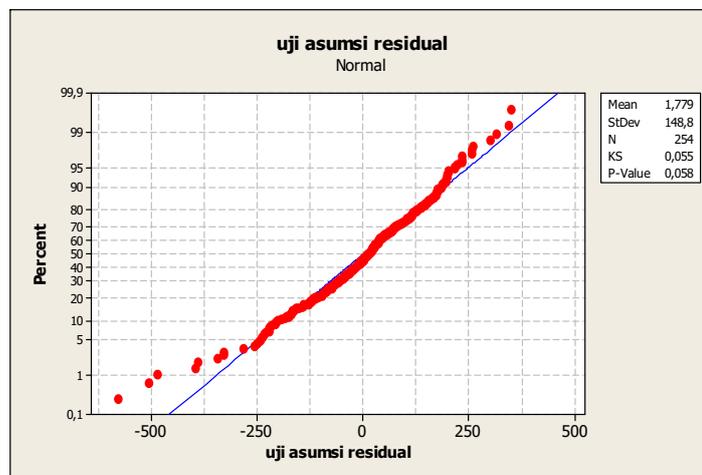
Pengujian asumsi independen bertujuan untuk mengetahui ada tidaknya korelasi antar residual. Salah satu cara untuk mendeteksi ada tidaknya korelasi antar residual adalah dengan uji Durbin-Watson. Berdasarkan output *matlab*, diperoleh  $d = 1.9517$ .

Tabel 10. Uji Durbin-Watson

$d$	$d_L$	$d_U$	$4 - d_U$	$4 - d_L$
1.9517	1,69312	1,89393	2,10607	2,30688

Berdasarkan tabel 10,  $d$  diantara  $d_U$  dan  $4 - d_U$  sehingga tidak terjadi autokorelasi.

Pengujian distribusi normal pada model regresi digunakan untuk menguji apakah nilai residual yang dihasilkan berdistribusi secara normal atau tidak. Model regresi yang bagus memiliki nilai residual yang berdistribusi normal. Metode uji normalitas menggunakan uji *Kolmogorov-Smirnov*.



Gambar 3. Uji Distribusi Normal Regresi *Spline* Terbobot

Dari gambar diperoleh informasi bahwa  $p - value = 0,058$ . Dengan menggunakan  $\alpha = 0,05$  keputusan yang diambil adalah gagal tolak  $H_0$  yang berarti residual berdistribusi normal.

#### Interpretasi Model

Interpretasi model dilakukan untuk parameter yang signifikan.

- Interpretasi terhadap variabel  $x_1$  adalah jika variabel  $x_2, x_3$  dan  $x_4$  konstan. Maka pengaruh hemoglobin ( $x_1$ ) terhadap kadar gula darah sewaktu ( $y$ ),

$$\hat{y}_i = 13,8465x_{1i} + 94,2316(x_{1i} - 10,8)_+ - 150,1617(x_{1i} - 11,7)_+ + 58,7381(x_{1i} - 13,4)_+$$

$$= \begin{cases} 13,8465x_{1i}, & x_{1i} < 10,8 \\ -1017,7013 + 108,0781x_{1i}, & 10,8 \leq x_{1i} < 11,7 \\ 739,1906 - 42,0836x_{1i}, & 11,7 \leq x_{1i} < 13,4 \\ -47,8999 + 16,6545x_{1i}, & x_{1i} \geq 13,4 \end{cases}$$

dari model di atas dapat diinterpretasikan pada saat kadar hemoglobin kurang dari 10,8 g/dL, apabila kadar hemoglobin naik sebesar 1g/dL, maka kadar GDS akan naik sebesar 13,8465 mg/dL. Kadar hemoglobin antara 10,8 g/dL sampai 11,7 g/dL apabila kadar hemoglobin naik sebesar 1 g/dL maka kadar GDS akan naik sebesar 108,0781mg/dL. Kadar hemoglobin antara 11,7 g/dL sampai 13,4 g/dL apabila kadar hemoglobin naik sebesar 1 g/dL maka kadar GDS akan turun sebesar 42,0836mg/dL. Apabila kadar hemoglobin lebih dari 13,4g/dL apabila kadar hemoglobin naik sebesar 1 g/dL maka kadar GDS akan naik sebesar 16,6545mg/dL

- b. Interpretasi terhadap variabel  $x_2$  adalah jika variabel  $x_1, x_3$  dan  $x_4$  konstan. Maka pengaruh leukosit ( $x_2$ ) terhadap kadar gula darah sewaktu ( $y$ ),

$$\hat{y}_i = -41,8248x_{2i} + 30,2921(x_{2i} - 3,8)_+ \\ = \begin{cases} -41,8248x_{2i} \\ -156,9348 + 30,2921x_{2i} \end{cases}$$

dari model di atas dapat diinterpretasikan pada saat kadar leukosit kurang dari 3,8 sel/ $\mu$ L, apabila kadar leukosit naik sebesar 1 sel/ $\mu$ L, maka kadar GDS akan turun sebesar 41,8248 mg/dL. Kadar leukosit antara 3,8 sel/ $\mu$ L sampai 6,7sel/ $\mu$ L apabila kadar leukosit naik sebesar 1 sel/ $\mu$ L maka kadar GDS akan naik sebesar 30,2921mg/dL.

- c. Interpretasi terhadap variabel  $x_3$  adalah jika variabel  $x_1, x_2$  dan  $x_4$  konstan. Maka pengaruh tekanan sistol ( $x_3$ ) terhadap kadar gula darah sewaktu ( $y$ ),

$$\hat{y}_i = -7,6663x_{3i}$$

dari model di atas dapat diinterpretasikan pada saat tekanan sistol kurang dari 120 mmHg, apabila tekanan sistol naik sebesar 1 mmHg, maka kadar GDS akan turun sebesar 7,6663 mg/dL.

- d. Interpretasi terhadap variabel  $x_4$  adalah jika variabel  $x_1, x_2$  dan  $x_3$  konstan. Maka pengaruh tekanan distol ( $x_4$ ) terhadap kadar gula darah sewaktu ( $y$ ),

$$\hat{y}_i = 7,2979x_{4i}$$

pada saat tekanan distol kurang dari 70 mmHg, apabila tekanan distol naik sebesar 1 mmHg, maka kadar GDS akan naik sebesar 7,2979 mg/dL.

### SIMPULAN

Model regresi *spline* linier terboboti dari faktor resiko penderita DM tipe II adalah sebagai berikut:

$$\hat{y}_i = 13,8465x_{1i} + 94,2316(x_{1i} - 10,8)_+ + -150,1617(x_{1i} - 11,7)_+ \\ + 58,7381(x_{1i} - 13,4)_+ - 41,8248x_{2i} + 30,2921(x_{2i} - 3,8)_+ \\ - 7,6663x_{3i} + 7,2979x_{4i}$$

Model regresi *spline* ini menghasilkan nilai GCV sebesar 24514,2137 dengan koefisien determinasi 24,98%. Variabel yang berpengaruh terhadap gula darah sewaktu adalah hemoglobin, leukosit, tekanan sistol, dan tekanan distol.

Interpretasi model regresi *Spline* terbaik adalah sebagai berikut:

1. Dengan asumsi kadar leukosit, tekanan sistol dan tekanan distol tetap  
Pada saat kadar hemoglobin kurang dari 11,7 g/dL dan lebih dari 13,4 g/dL apabila kadar hemoglobin naik maka kadar GDS akan naik. Kadar hemoglobin antara 11,7 g/dL sampai 13,4 g/dL apabila kadar hemoglobin naik maka presentase GDS akan turun.
2. Dengan asumsi kadar hemoglobin, tekanan sistol, dan tekanan distol tetap

Pada saat kadar leukosit kurang dari  $3,8 \text{ sel}/\mu\text{L}$ , apabila kadar leukosit naik, maka kadar GDS akan turun. Kadar leukosit antara  $3,8 \text{ sel}/\mu\text{L}$  sampai  $6,7 \text{ sel}/\mu\text{L}$  apabila kadarleukosit naik maka kadar GDS akan naik.

3. Dengan asumsi kadar hemoglobin, kadar leukosit, dan tekanan ditol tetap Pada saat tekanan sistol kurang dari  $120 \text{ mmHg}$  ,apabila presentase tekanan sistol naik, maka kadar GDS akan turun.
4. Dengan asumsi kadar hemoglobin, kadar leukosit, dan tekanan sistol tetap Pada saat tekanan distol kurang dari  $70 \text{ mmHg}$  ,apabila tekanan distol naik, maka kadar GDS akan naik.

#### DAFTAR PUSTAKA

- [1] P. E. Indonesia, “Pengelolaan dan pencegahan diabetes melitus tipe 2 di Indonesia,” *Pb. Perkeni*, 2015.
- [2] C. EJ, “Buku Saku Patofisiologi, Ed ke-3,” *Jakarta Penerbit Buku Kedokt. EGC*, pp. 477–481, 2007.
- [3] I. P. Merdekawati and I. N. Budiantara, “Pemodelan Regresi Spline Truncated Multivariabel pada Faktor-Faktor yang Mempengaruhi Kemiskinan di Kabupaten/Kota Provinsi Jawa Tengah,” *J. Sains dan Seni ITS*, vol. 2, no. 1, pp. D19–D24, 2013.
- [4] R. L. Eubank, *Nonparametric regression and spline smoothing*. CRC press, 1999.
- [5] S. C. Smeltzer and B. G. Bare, “Buku ajar keperawatan medikal bedah,” *Jakarta Egc*, vol. 1223, p. 21, 2002.
- [6] H. Wijayakusuma, *Bebas Diabetes ala Hembing*. Niaga Swadaya, 2008.
- [7] C. Arthur, M. D. Guyton, and J. E. Hall, “Textbook of medical physiology,” *Mississippi Guyt. Hall*, 2006.
- [8] Z. Effendi, “Peranan Leukosit sebagai Anti Inflamasi Alergik dalam Tubuh,” 2003.
- [9] D. Garnita, “Faktor Risiko Diabetes Melitus di Indonesia (Analisis Data Sakerti 2007),” *Fak. Kesehat. Masyarkat Univ. Indones. Depok*, 2012.
- [10] M. F. Bintariningrum and I. N. Budiantara, “Pemodelan regresi nonparametrik spline truncated dan aplikasinya pada Angka Kelahiran Kasar di Surabaya,” *J. Sains dan Seni ITS*, vol. 3, no. 1, pp. D7–D12, 2014.
- [11] A. Tripena, “Penentuan Model Regresi Spline Terbaik,” in *Prosiding Seminar Nasional Statistika Universitas Diponegoro 2011*, 2011, pp. 92–102.