

PEMODELAN FAKTOR-FAKTOR YANG MEMPENGARUHI DIABETES MELITUS TIPE 2 DI INDONESIA TAHUN 2023 MENGGUNAKAN METODE REGRESI NONPARAMETRIK *SPLINE TRUNCATED*

Katarina Rosa Leonida

Mahasiswa Program Studi Statistika, Fakultas Teknik dan Sains, Universitas PGRI Adi Buana Surabaya

e-mail : rosaleonydha@gmail.com

Artanti Indrasetianingsih*

Dosen Program Studi Statistika, Fakultas Teknik dan Sains, Universitas PGRI Adi Buana Surabaya

e-mail : artanti.indra@unipasby.ac.id

Gangga Anuraga

Dosen Program Studi Statistika, Fakultas Teknik dan Sains, Universitas PGRI Adi Buana Surabaya

e-mail : g.anuraga@unipasby.ac.id

Muhammad Athoillah

Dosen Program Studi Statistika, Fakultas Teknik dan Sains, Universitas PGRI Adi Buana Surabaya

e-mail : athoillah@unipasby.ac.id

Abstrak

Penelitian ini bertujuan untuk menganalisis faktor-faktor yang mempengaruhi Diabetes melitus tipe 2 (DM2) di Indonesia tahun 2023 dengan menggunakan metode regresi nonparametrik *spline truncated*. Metode ini dipilih karena fleksibilitasnya dalam menganalisis pola yang tidak linier dalam data melalui penggunaan fungsi *spline* dan titik knot. Pemilihan titik knot optimal dapat dilakukan dengan menggunakan metode *Generalized Cross Validation* (GCV). Data sekunder dalam penelitian ini diperoleh dari Badan Pusat Statistik (BPS) dan Survey Kesehatan Indonesia yang mencakup 38 provinsi di Indonesia. Lima variabel independen yang dianalisis meliputi proporsi konsumsi makanan manis, proporsi aktivitas fisik, proporsi konsumsi rokok tembakau, prevalensi penderita obesitas dan proporsi konsumsi makanan olahan. Hasil penelitian menunjukkan bahwa model regresi nonparametrik *spline truncated* terbaik adalah menggunakan tiga titik knot. Terdapat empat variabel dalam penelitian yang berpengaruh signifikan yaitu variabel proporsi konsumsi makanan manis, proporsi konsumsi rokok tembakau, prevalensi penderita obesitas, dan proporsi konsumsi makanan olahan.

Kata Kunci: Diabetes Melitus Tipe 2, Regresi Nonparametrik *Spline Truncated*, Titik Knot, Indonesia

Abstract

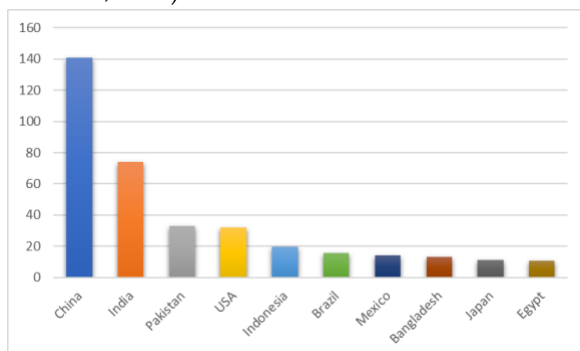
This study aimed to analyze the factors affecting type 2 diabetes mellitus (DM2) in Indonesia in 2023 using the *spline truncated nonparametric regression method*. This method was chosen because of its flexibility in analyzing non-linear patterns in the data through the use of *spline functions* and *knots*. Selection of optimal knot points can be done using the *Generalized Cross Validation* (GCV) method. Secondary data in this study were obtained from the Central Bureau of Statistics (BPS) and the Indonesian Health Survey covering 38 provinces in Indonesia. Five independent variables were analyzed, including the proportion of sweet food consumption, the proportion of physical activity, the proportion of tobacco cigarette consumption, the prevalence of obese people and the proportion of processed food consumption. The results showed that the best truncated *spline nonparametric regression model* was using three knot points. The study identified four significant variables: proportion of sweet food consumption, proportion of tobacco cigarette consumption, prevalence of obese people, and proportion of processed food consumption.

Keywords: Diabetes Mellitus Type 2, Truncated *Spline Nonparametric Regression*, Knot Points, Indonesia

PENDAHULUAN

Kesehatan merupakan aspek penting dalam kehidupan, yang berperan dalam meningkatkan kesejahteraan dan produktivitas masyarakat. Namun, berbagai tantangan dalam menjaga kesehatan seringkali muncul akibat gaya hidup yang kurang sehat. Salah satu dampak dari pola hidup yang tidak sehat adalah meningkatnya prevalensi penyakit tidak menular, termasuk diabetes melitus. Diabetes melitus (DM) adalah penyakit metabolik yang ditandai dengan peningkatan kadar glukosa darah (hiperglikemia) akibat gangguan produksi atau efektivitas insulin.

Terdapat tiga jenis klasifikasi atau tipe diabetes yaitu diabetes melitus tipe 1, diabetes melitus tipe 2, dan diabetes melitus gestasional (Delfina et al., 2021). Dari ketiga jenis klasifikasi tersebut, kurang lebih 90-95% diabetes melitus tipe 2 menyumbang banyak dari total kasus diabetes global, sehingga menjadikannya sebagai penyakit penyebab kematian ke-3 di dunia (PERKENI, 2021). Diabetes Melitus tipe 2 yaitu penyakit diabetes tidak tergantung insulin yang terjadi karena penurunan sensitivitas insulin (resistensi insulin) akibat berkurangnya jumlah insulin yang diproduksi oleh pankreas. Data dari *International Diabetes Federation (IDF) Diabetes Atlas* menunjukkan prevalensi diabetes global pada usia 20-79 tahun diperkirakan mencapai 10,5% (536,6 juta orang) pada tahun 2021 dan diprediksi meningkat menjadi 12,2% (783,2 juta orang) pada tahun 2045 (Sun et al., 2021).



Gambar 1. Negara dengan Jumlah Kasus Diabetes Tertinggi di Dunia Tahun 2021

Dari gambar di atas diketahui bahwa, Indonesia berada di peringkat ke-5 dunia dalam jumlah pengidap diabetes sebanyak 19,47 juta penderita pada tahun 2021 dan angka ini diprediksi meningkat menjadi 28,6 juta pada tahun 2045. Peningkatan kasus diabetes melitus tipe 2 di Indonesia menunjukkan

perlu kesadaran masyarakat terhadap pentingnya gaya hidup sehat. Namun, masih banyak masyarakat yang kurang peduli terhadap faktor-faktor risiko yang dapat dicegah seperti pola makan yang kurang sehat, kurang aktivitas fisik, gangguan lemak darah, obesitas dan tekanan darah tinggi (Association, 2020). Melihat tingginya angka kejadian diabetes melitus tipe 2, upaya preventif harus segera ditingkatkan, terutama dalam meningkatkan pemahaman masyarakat mengenai pola hidup sehat.

Penelitian sebelumnya mengenai diabetes melitus tipe 2 telah dilakukan oleh beberapa peneliti, salah satunya penelitian yang dilakukan oleh (Safire & Purhadi, 2023) tentang “Pemodelan Faktor-Faktor yang mempengaruhi Jumlah Kasus Diabetes Melitus di Jawa Timur Menggunakan GWGPR dan GWNBR”. Dari penelitian tersebut diperoleh faktor-faktor yang berpengaruh signifikan terhadap jumlah kasus diabetes melitus antara lain rasio pendidikan SMA ke-atas, proporsi aktivitas fisik penduduk, persentase konsumsi rokok tembakau, persentase penderita hipertensi, dan proporsi konsumsi minuman manis. Penelitian lain dilakukan oleh (Miranda & Sari, 2023) tentang “Faktor Risiko Kejadian Diabetes Melitus (DM) Tipe 2 Di Poliklinik Penyakit Dalam RS Kota Bengkulu”. Hasil penelitian menunjukkan bahwa variabel yang berhubungan dengan kejadian Diabetes Melitus tipe 2 adalah usia, riwayat keluarga, IMT, aktivitas fisik, pola makan tinggi gula, dan pola makan rendah serat. Variabel dominan yang berhubungan dengan kejadian Diabetes Melitus tipe 2 adalah aktivitas fisik. Orang yang aktivitas fisiknya kurang berpeluang tinggi menderita diabetes melitus tipe 2 daripada orang yang aktivitas fisiknya cukup. Pada konteks data longitudinal, penelitian oleh (Putri, 2018) menerapkan “Regresi Nonparametrik *Spline Truncated* untuk Memodelkan Persentase Peserta KB Baru Pasca Persalinan Dan Pasca Keguguran di Jawa Timur”. Model terbaik ditemukan menggunakan kombinasi titik knot yang optimal menghasilkan koefisien determinasi sebesar 90,71% dan menunjukkan bahwa semua variabel prediktor berpengaruh signifikan. (Ramadhan et al., 2024) melakukan penelitian tentang penerapan “Regresi Nonparametrik *Spline Truncated* pada Indeks Pembangunan Gender (IPG) di Sumatera Utara”. Penelitian ini mengidentifikasi faktor-faktor yang

signifikan mempengaruhi IPG dengan pemilihan titik knot optimal menggunakan metode *Generalized Cross Validation* (GCV), dengan nilai GCV minimum sebesar 5,64129. Hasil menunjukkan bahwa *spline truncated* efektif dalam menangkap fluktuasi yang tidak teratur.

Dalam penelitian ini, faktor-faktor yang diduga berkontribusi signifikan terhadap kejadian diabetes melitus tipe 2 mencakup konsumsi makanan manis, kebiasaan merokok, kurangnya aktivitas fisik, konsumsi makanan olahan, dan prevalensi penderita obesitas. Pemilihan variabel ini didasarkan pada hasil penelitian sebelumnya dalam menjelaskan faktor-faktor yang berpengaruh terhadap diabetes melitus tipe 2. Konsumsi makanan manis dan makanan olahan berhubungan dengan peningkatan kadar glukosa dalam darah. Kebiasaan merokok dikaitkan dengan resistensi insulin dan peningkatan risiko komplikasi diabetes. Kurangnya aktivitas fisik berkontribusi pada peningkatan IMT dan resistensi insulin, sementara obesitas merupakan faktor risiko utama yang telah terbukti meningkatkan kemungkinan seseorang terkena diabetes melitus tipe 2.

Dalam upaya memahami faktor-faktor yang mempengaruhi diabetes melitus tipe 2, analisis statistik menjadi alat yang sangat penting dalam mengidentifikasi hubungan antara variabel independen terhadap variabel dependen yaitu persentase diabetes melitus tipe 2. Namun, hubungan antara variabel-variabel tersebut seringkali tidak linier dan kompleks, yang berarti perubahan pada variabel independen tidak selalu menghasilkan perubahan yang proporsional pada variabel dependen. Pendekatan seperti regresi linier mengasumsikan bahwa hubungan antar variabel bersifat linier. Namun, asumsi ini seringkali tidak sesuai dalam kasus diabetes melitus tipe 2 karena adanya pola hubungan yang berbeda dan bervariasi pada berbagai rentang data. Untuk mengatasi keterbatasan ini, diperlukan metode statistik yang fleksibel dan adaptif untuk menganalisis data dengan pola yang tidak linier seperti regresi nonparametrik.

Regresi nonparametrik yang cukup populer adalah *spline*. Regresi nonparametrik *Spline* merupakan fungsi potongan polinomial yang terdiri dari segmen-segmen polinomial yang saling terhubung di titik-titik tertentu yang disebut knots,

Spline memungkinkan model untuk mengadaptasi berbagai bentuk data dengan lebih bebas dan dapat menghasilkan estimasi yang halus di seluruh rentang data. Namun *spline* dapat menyebabkan estimasi yang tidak akurat jika terdapat outlier atau ekstrem pada data. Sebagai alternatif untuk menghindari hal tersebut, *spline truncated* dapat digunakan untuk mengatasi outlier atau ekstrem pada data (Wongkar & Ruliana, 2023).

Regresi nonparametrik *spline truncated* adalah metode analisis regresi yang efektif dalam mengestimasi data tanpa pola tertentu dan cenderung dapat menyesuaikan estimasi berdasarkan pola yang muncul dari data tersebut (Budiantara, 2011). Keunggulan utama metode ini adalah fleksibilitasnya yang tinggi dalam membentuk kurva regresi, mempunyai sifat tersegmen, dan dapat menyesuaikan diri terhadap karakteristik lokal suatu data.

KAJIAN TEORI

REGRESI NONPARAMETRIK *SPLINE TRUNCATED*

Regresi nonparametrik merupakan suatu metode untuk mengetahui pola hubungan antara variabel independen dan variabel dependen yang tidak diketahui bentuk fungsinya. Pendekatan ini memungkinkan model untuk menangkap hubungan yang lebih kompleks dan tidak terstruktur antara variabel independen dan dependen. Misalnya, jika hubungan antara dua variabel bersifat nonlinear atau memiliki pola yang tidak biasa, regresi nonparametrik dapat menyesuaikan diri dengan pola tersebut. Oleh karena itu pendekatan model regresi nonparametrik sangat fleksibel dalam memodelkan pola data (Eubank, 1999). Secara umum model regresi nonparametrik dapat ditulis sebagai berikut :

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

Fungsi regresi nonparametrik *spline truncated* diperoleh dari penjumlahan antara fungsi polinomial dengan fungsi truncated. Fungsi *spline truncated* dengan titik-titik knot dapat dituliskan menjadi persamaan sebagai berikut (Budiantara, 2019).

$$f(x) = \sum_{j=0}^p \beta_j x_i^j + \sum_{k=1}^r \beta_{p+k} (x_i - K_k)_+^p \quad (2)$$

Apabila persamaan (2) disubstitusikan ke persamaan (1) maka akan diperoleh persamaan model regresi nonparametrik *spline truncated* sebagai berikut :

$$y_i = \sum_{j=0}^p \beta_j x_i^j + \sum_{k=1}^r \beta_{p+k} (x_i - K_k)_+^p + \varepsilon_i, i = 1, 2, \dots, n \quad (3)$$

dimana :

- β_j : parameter model polinomial
- $j = 1, 2, \dots, p$: derajat polinomial
- β_{p+k} : parameter pada komponen *truncated*
- $k = 1, 2, \dots, r$: banyaknya titik knot
- x_i : variabel independen ke-i
- ε_i : error untuk observasi ke-i
- $(x_i - K_k)_+^p$: fungsi *truncated*

ESTIMASI PARAMETER REGRESI

NONPARAMETRIK SPLINE TRUNCATED

Salah satu metode yang digunakan dalam mengestimasi parameter model dalam regresi nonparametrik spline truncated adalah metode *Ordinary Least Square* (OLS). Metode OLS bertujuan untuk menemukan penduga parameter regresi dengan meminimalkan jumlah kuadrat residual. Dalam notasi matriks penulisan regresi nonparametrik spline dapat dinyatakan pada persamaan (4) (Kurniawati & Budiantara, 2019).

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

dimana,

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} \cdots x_{1q} & (x_{11} - K_1)_+^p & \cdots & (x_{1q} - K_{qr})_+^p \\ 1 & x_{21} \cdots x_{2q} & (x_{21} - K_1)_+^p & \cdots & (x_{2q} - K_{qr})_+^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} \cdots x_{nq} & (x_{n1} - K_1)_+^p & \cdots & (x_{nq} - K_{qr})_+^p \end{bmatrix}, \quad (5)$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_q \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Berdasarkan persamaan (5), persamaan residual dapat dituliskan dalam bentuk persamaan di bawah.

$$\varepsilon = y - X\beta \quad (6)$$

Jumlah kuadrat residual dalam bentuk matriks dapat ditulis sebagai berikut

$$\begin{aligned} \sum_{i=1}^n \varepsilon_i^2 &= \varepsilon' \varepsilon \\ &= (y - X\beta)'(y - X\beta) \\ &= y'y - y'X\beta - \beta'X'y + \beta'X'X\beta \\ &= y'y - 2\beta'X'y + \beta'X'X\beta \end{aligned} \quad (7)$$

Untuk meminimumkan nilai $\varepsilon' \varepsilon$, maka turunan pertama terhadap β harus sama dengan nol.

$$\frac{\partial(\varepsilon' \varepsilon)}{\partial \beta} = 0 \quad (8)$$

Dari turunan pertama pada persamaan (8) diperoleh nilai estimasi parameter β yang ditunjukkan pada persamaan berikut.

$$\begin{aligned} -2X'y + 2X'X\hat{\beta} &= 0 \\ X'X\hat{\beta} &= X'y \\ (X'X)^{-1}(X'X)\hat{\beta} &= (X'X)^{-1}X'y \\ \hat{\beta} &= (X'X)^{-1}X'y \end{aligned} \quad (9)$$

PEMILIHAN TITIK KNOT OPTIMAL

Knot adalah suatu titik fokus dalam fungsi *spline* untuk menentukan di mana perubahan bentuk fungsi *spline* terjadi. Titik-titik ini berfungsi sebagai titik transisi di mana fungsi *spline* dapat berubah. Hal ini memungkinkan *spline* untuk menangkap pola hubungan yang lebih kompleks, sehingga model bisa lebih fleksibel dalam menyesuaikan diri terhadap perubahan pola dalam data dan prediksinya bisa lebih tepat atau akurat. Dalam penelitian ini, titik knot diperoleh dengan cara membagi rentang data dari masing-masing variabel independen ke dalam sejumlah titik secara merata yaitu lima puluh titik. Dari lima puluh titik tersebut, titik paling awal dan akhir dihilangkan untuk menghindari efek ekstrem pada data. Proses ini menghasilkan sekumpulan kandidat knot yang kemudian diuji melalui proses *grid search*. *Grid search* adalah metode pencarian sistematis untuk menemukan titik knot terbaik dengan cara mencoba berbagai kemungkinan titik knot secara menyeluruh. Setelah titik knot ditentukan, selanjutnya adalah pemilihan titik knot optimal dengan *Generalized Cross Validation* (GCV). Berikut merupakan fungsi GCV pada regresi nonparametrik secara umum dapat dituliskan sebagai berikut (Eubank, 1999).

$$GCV(K_1, K_2, \dots, K_r) = \frac{n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\left[n^{-1} \text{trace}(I - A) \right]^2} \quad (10)$$

dimana :

- K_1, K_2, \dots, K_r : titik knot awal hingga titik knot ke-r
- $GCV(K_1, K_2, \dots, K_r)$: vektor berisikan nilai GCV dari titik knot

\hat{y}_i : nilai prediksi dari variabel dependen
 $(\mathbf{I} - \mathbf{A})$: \mathbf{I} merupakan matriks identitas dan serta $\mathbf{A} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$
 n : banyaknya observasi

PENGUJIAN PARAMETER MODEL REGRESI

Pengujian parameter model regresi dibagi menjadi dua tahap yaitu pengujian parameter secara serentak dan pengujian parameter secara parsial. Uji serentak bertujuan untuk melihat secara serentak pengaruh variabel independen terhadap variabel dependen.

Hipotesis :

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{p+k} = 0$$

$$H_1 : \text{minimal ada satu } \beta_j \neq 0, j = 1, 2, \dots, p + k$$

Dengan statistik uji dinyatakan pada persamaan berikut.

$$F_{hitung} = \frac{MS_{regresi}}{MS_{error}} \tag{11}$$

Kriteria pengambilan keputusan adalah tolak H_0 apabila $F_{hitung} > F_{\alpha;(p+k, n-(p+k)-1)}$.

Uji parsial digunakan untuk melihat pengaruh masing-masing variabel independen terhadap variabel dependen.

Hipotesis :

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, \dots, p + k$$

Statistik uji yang digunakan dalam pengujian ini dituliskan dalam persamaan sebagai berikut (Draper & Smith, 1998).

$$t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \tag{12}$$

Kriteria pengambilan keputusannya adalah tolak H_0 jika $t_{hitung} > t_{(\alpha/2; n-(p+k)-1)}$ atau $p\text{-value} < \alpha$ yang berarti ada pengaruh antara variabel independen terhadap variabel dependen.

PENGUJIAN ASUMSI RESIDUAL

Terdapat beberapa pengujian asumsi yang harus dipenuhi dalam melakukan analisis menggunakan regresi nonparametrik *spline truncated* diantaranya

identik, independen, dan berdistribusi normal (IIDN) (Alamsyah et al., 2022).

● Uji Identik

Uji asumsi identik dapat dilakukan menggunakan Uji Glejser. Uji Glejser dilakukan dengan meregresikan absolut dari residual dengan variabel independennya dengan hipotesis sebagai berikut.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{minimal terdapat satu } \sigma_i^2 \neq \sigma^2, i = 1, 2, \dots, n$$

Statistik uji :

$$F_{hitung} = \frac{MSR}{MSE} = \frac{\left[\sum_{i=1}^n (|\varepsilon_i| - |\bar{\varepsilon}|)^2 \right] / (v-1)}{\left[\sum_{i=1}^n (|\varepsilon_i| - |\varepsilon_i|)^2 \right] / (n-v)} \tag{13}$$

Jika ditetapkan tingkat signifikan sebesar α , maka tolak H_0 apabila $F_{hitung} > F_{\alpha;(v, n-v-1)}$, dimana, $v=p+r$ dan v menunjukkan banyaknya parameter model glejser. Apabila H_0 ditolak, berarti residual tidak identik atau terjadi heteroskedastisitas.

● Uji Independen

Uji asumsi independen bertujuan untuk mengetahui apakah ada korelasi yang signifikan antara dua atau lebih variabel independen dalam model regresi. Salah satu cara untuk mendeteksi apakah terdapat korelasi antar variabel residual adalah dengan menggunakan uji Durbin-Watson. Hipotesis yang digunakan adalah sebagai berikut :

$$H_0 : \rho = 0 \text{ (tidak terdapat autokorelasi residual)}$$

$$H_1 : \rho \neq 0 \text{ (terdapat autokorelasi residual)}$$

Statistik uji :

$$dW = \frac{\sum_{i=1}^n (\varepsilon_i - \varepsilon_{i-n})^2}{\sum_{i=1}^n \varepsilon_i^2} \tag{14}$$

Kriteria pengambilan keputusan gagal tolak H_0 jika $dW < dL$, maka terjadi autokorelasi antar variabel independent, jika $dL < dW < dU$ maka keputusan tidak pasti, dan jika $dU < dW < 4-dU$ maka tidak terjadi autokorelasi antar variabel.

● Uji Normalitas

Uji normalitas adalah sebuah uji yang dilakukan dengan tujuan untuk menentukan data yang dikumpulkan berdistribusi normal atau tidak. Pengujian normalitas dapat dilakukan dengan menggunakan uji Kolmogorov-Smirnov dengan hipotesis sebagai berikut.

$H_0 : F_n(\varepsilon) = F_0(\varepsilon)$ (Residual berdistribusi normal)

$H_1 : F_n(\varepsilon) \neq F_0(\varepsilon)$ (Residual tidak berdistribusi normal)

Statistik uji :

$$D = \text{maksimal} |F_n(\varepsilon) - F_0(\varepsilon)| \quad (15)$$

Kriteria pengambilan keputusan dapat ditentukan dengan melihat $|D| < D_{\alpha,n}$ atau $p\text{-value} > \alpha$ maka gagal tolak H_0 . Dengan demikian residual data berdistribusi normal dan asumsi terpenuhi.

METODE

DATA PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari data Laporan Survey Kesehatan Indonesia (SKI) tahun 2023, Dinas Kesehatan Indonesia tahun 2023, dan Badan Pusat Statistik (BPS) Indonesia tahun 2023. Unit observasi pada penelitian yaitu 38 Provinsi yang ada di Indonesia. Berikut rincian variabel penelitian yang digunakan.

Tabel 1. Variabel Penelitian

Variabel	Keterangan	Skala Data
Y	Persentase Diabetes Melitus Tipe 2	Rasio
X_1	Proporsi Konsumsi Makanan Manis	Rasio
X_2	Proporsi Aktivitas Fisik	Rasio
X_3	Proporsi Konsumsi Rokok Tembakau	Rasio
X_4	Prevalensi Penderita Obesitas	Rasio
X_5	Proporsi Konsumsi Makanan Olahan	Rasio

TAHAPAN PENELITIAN

Berikut adalah langkah-langkah yang dilakukan untuk menganalisis data dalam penelitian ini adalah sebagai berikut :

1. Mengumpulkan data
2. Mendeskripsikan karakteristik data persentase diabetes melitus tipe 2 di Indonesia beserta faktor-faktor yang diduga mempengaruhinya.
3. Mengidentifikasi pola hubungan antara variabel dependen dengan masing-masing variabel independen dengan menggunakan *scatterplot*.
4. Memodelkan persentase diabetes melitus tipe 2 menggunakan metode Regresi Nonparametrik *Spline Truncated*.

5. Memilih titik knot optimal berdasarkan nilai GCV yang paling minimum.
6. Mendapatkan model regresi nonparametrik *spline truncated* terbaik dengan titik knot optimal.
7. Menguji signifikansi parameter model yang telah didapatkan secara serentak dan parsial.
8. Melakukan uji asumsi residual identik, independen, dan berdistribusi normal (IIDN).
9. Menghitung nilai koefisien determinasi R^2 .
10. Menginterpretasikan model
11. Membuat kesimpulan.

HASIL DAN PEMBAHASAN

STATISTIK DESKRIPTIF

Analisis statistika deskriptif digunakan untuk mengetahui karakteristik dari diabetes melitus tipe 2 serta faktor-faktor yang mempengaruhinya. Berikut disajikan karakteristik dari kelima variabel yang diduga mempengaruhi diabetes melitus tipe 2 di Indonesia.

Tabel 2. Statistika Deskriptif Variabel Penelitian

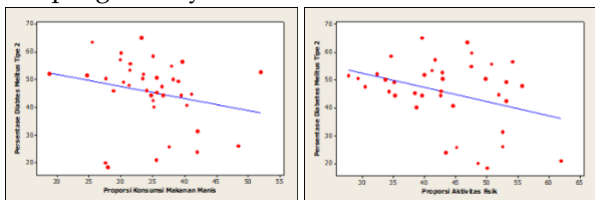
Variabel	Min	Max	Mean	Standar Deviasi
Y	18,40	65,10	45,57	12,06
X_1	18,80	52,00	34,70	6,36
X_2	27,80	62,00	43,50	8,19
X_3	0,20	10,90	2,19	3,28
X_4	13,30	31,80	22,97	4,48
X_5	24,50	58,70	43,49	8,92

Tabel 2. menunjukkan bahwa rata-rata dari variabel diabetes melitus tipe 2 (Y) di Indonesia sebesar 45,57 persen dengan standar deviasi sebesar 12,06. Provinsi Kalimantan Barat merupakan provinsi yang memiliki persentase terbesar terdiagnosis diabetes melitus tipe 2 yaitu 65,10 persen, sedangkan Provinsi Papua Selatan merupakan provinsi yang memiliki persentase terkecil terdiagnosis diabetes melitus tipe 2 yaitu 18,40 persen. Variabel proporsi konsumsi makanan manis (X_1) memiliki rata-rata sebesar 34,70 persen dengan nilai standar deviasi sebesar 6,36. Provinsi dengan proporsi konsumsi makanan manis tertinggi adalah Kalimantan selatan yaitu sebesar 52,0 persen, sedangkan provinsi dengan proporsi konsumsi makanan manis terendah adalah Bali yaitu sebesar 18,80 persen. Variabel proporsi aktivitas fisik (X_2) memiliki rata-rata sebesar 43,50 dengan standar

deviasi sebesar 8,19. Provinsi dengan proporsi aktivitas fisik tertinggi adalah Papua Tengah yaitu sebesar 62,00 persen, sedangkan proporsi aktivitas fisik terendah adalah Provinsi Jawa Timur yaitu sebesar 27,80 persen. Proporsi konsumsi rokok tembakau (X_3) memiliki rata-rata sebesar 2,19 persen dengan standar deviasi sebesar 3,28. Provinsi dengan proporsi konsumsi rokok tembakau tertinggi adalah Papua Barat yaitu sebesar 10,90 persen, sedangkan provinsi dengan proporsi konsumsi rokok tembakau terendah adalah Bangka Belitung dan DKI Jakarta yaitu sebesar 0,2 persen. Prevalensi penderita obesitas (X_4) memiliki rata-rata sebesar 22,97 persen dengan standar deviasi 4,48. Prevalensi penderita obesitas tertinggi terdapat di Provinsi DKI Jakarta yaitu sebesar 31,80 persen, sedangkan prevalensi penderita obesitas terendah terdapat di Provinsi NTT yaitu sebesar 13,30 persen. Proporsi konsumsi makanan olahan (X_5) memiliki rata-rata sebesar 43,49 persen dengan standar deviasi 8,92. Provinsi dengan proporsi konsumsi makanan olahan tertinggi adalah Kalimantan Tengah yaitu sebesar 58,70 persen, sedangkan proporsi konsumsi makanan olahan terendah adalah Provinsi Sulawesi Gorontalo yaitu sebesar 24,50 persen.

ANALISIS POLA HUBUNGAN

Identifikasi pola hubungan antar variabel dapat dilakukan menggunakan scatterplot. Hasil dari scatterplot dapat menunjukkan apakah pola hubungan antara variabel dependen dan variabel independen memiliki pola tertentu atau tidak. Berikut merupakan scatterplot antara persentase diabetes melitus tipe 2 dengan faktor-faktor yang mempengaruhinya.

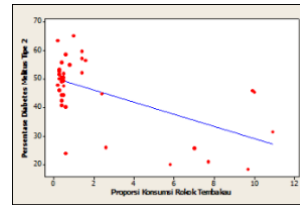


Gambar 1.

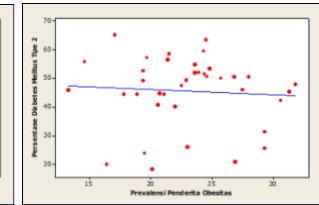
Gambar 2.

Gambar 1. menunjukkan pola hubungan antara variabel persentase diabetes melitus tipe 2 dan proporsi konsumsi makanan manis dan Gambar 2. menunjukkan pola hubungan antara persentase diabetes melitus tipe 2 dan proporsi aktivitas fisik. Terlihat pada kedua gambar tersebut bahwa, penyebaran titik-titik pada grafik cenderung

menyebar dari kiri atas ke kanan bawah dan kemiringan garis regresi ke bawah, yang menunjukkan adanya pola hubungan yang negatif. Secara umum, jika proporsi konsumsi makanan manis dan proporsi aktivitas fisik semakin tinggi, maka persentase diabetes melitus tipe 2 cenderung menurun.

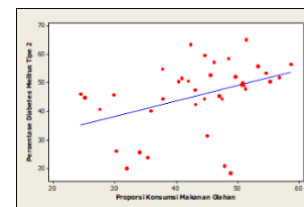


Gambar 3.



Gambar 4.

Gambar 3. menunjukkan pola hubungan antara variabel persentase diabetes melitus tipe 2 dan proporsi konsumsi rokok tembakau dan Gambar 2. menunjukkan pola hubungan antara persentase diabetes melitus tipe 2 dan prevalensi penderita obesitas. Pada Gambar 3. dapat dilihat bahwa adanya kecenderungan hubungan negatif yang dapat dilihat kemiringan pada garis biru regresi ke bawah. Artinya, secara umum, semakin tinggi proporsi konsumsi rokok tembakau, maka persentase penderita diabetes melitus tipe 2 cenderung menurun. Pada Gambar 4. terlihat bahwa garis regresi cenderung datar dengan sedikit kemiringan negatif, yang berarti bahwa secara umum terdapat hubungan negatif yang sangat lemah antara prevalensi obesitas dan persentase diabetes melitus tipe 2. Artinya, ketika prevalensi penderita obesitas meningkat, maka persentase diabetes melitus tipe 2 cenderung sedikit menurun.



Gambar 5.

Gambar 5. menunjukkan pola hubungan antara persentase diabetes melitus tipe 2 dan proporsi konsumsi makanan olahan. Dari gambar terlihat bahwa garis biru regresi memiliki kemiringan ke atas, yang berarti terdapat hubungan positif antara konsumsi makanan olahan dan persentase penderita diabetes melitus Tipe 2. Secara umum, semakin tinggi proporsi konsumsi makanan olahan di suatu wilayah, maka semakin tinggi pula persentase penderita diabetes melitus Tipe 2.

PEMILIHAN TITIK KNOT OPTIMUM

Berikut merupakan pemilihan titik knot optimal dengan satu knot, dua knot, dan tiga knot.

Tabel 3. Nilai GCV dengan satu titik knot

X_1	X_2	X_3	X_4	X_5	GCV
22,865	31,987	1,5102	15,565	28,687	89,330
22,187	31,289	1,2918	15,187	27,989	89,404
23,542	32,685	1,7285	15,942	29,385	91,449
21,510	30,591	1,0734	14,810	27,291	92,676
24,220	33,383	1,946	16,320	30,083	93,943

Berdasarkan Tabel 3. didapatkan nilai GCV paling minimum untuk model regresi nonparametrik *spline truncated* dengan satu titik knot adalah sebesar 89,330. Titik knot optimal untuk variabel proporsi konsumsi makanan manis (X_1) berada pada titik knot 22,863, variabel proporsi aktivitas fisik (X_2) berada pada titik knot 31,987, variabel proporsi rokok tembakau (X_3) berada pada titik knot 1,510, variabel prevalensi penderita obesitas (X_4) berada pada titik knot 15,565, dan variabel proporsi konsumsi makanan olahan (X_5) berada pada titik knot 28,687. Berikut merupakan hasil pemodelan regresi nonparametrik *spline truncated* menggunakan satu titik knot dengan GCV paling minimum.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - 22,86351)_+ + \hat{\beta}_3 x_2 + \hat{\beta}_4 (x_2 - 31,98776)_+ + \hat{\beta}_5 x_3 + \hat{\beta}_6 (x_3 - 1,51020)_+ + \hat{\beta}_7 x_4 + \hat{\beta}_8 (x_4 - 15,56531)_+ + \hat{\beta}_9 x_5 + \hat{\beta}_{10} (x_5 - 28,68776)_+$$

Estimasi paramater model regresi nonparametrik *spline truncated* dengan dua titik knot dapat dilihat pada Tabel 4 di bawah.

Tabel 4. Nilai GCV dengan Dua Titik Knot

X_1	X_2	X_3	X_4	X_5	GCV
24,897	34,081	2,1653	16,697	30,781	61,657
26,930	36,175	2,8204	17,830	32,875	
25,575	34,779	2,3836	17,075	31,479	64,171
26,930	36,175	2,8204	15,187	32,875	
22,187	31,289	1,2918	16,320	27,989	65,423
28,285	37,571	3,257	18,208	34,271	
24,220	33,383	1,946	16,320	30,083	65,679
27,608	36,873	3,038	18,585	33,573	
24,220	33,383	1,946	15,565	30,083	66,854
28,285	37,571	3,257	18,585	34,271	

Berdasarkan Tabel 4 didapatkan nilai GCV paling minimum unuk model regresi nonparametrik *spline truncated* dengan dua titik knot adalah 61,657. Titik knot optimal untuk masing-masing variabel adalah sebagai berikut.

- a) Proporsi konsumsi makanan manis (X_1)

$$K_1 = 24,89796 \text{ dan } K_2 = 26,93061$$

- b) Proporsi aktivitas fisik (X_2)

$$K_1 = 34,08163 \text{ dan } K_2 = 36,17551$$

- c) Proporsi konsumsi rokok tembakau (X_3)

$$K_1 = 2,165306 \text{ dan } K_2 = 2,820408$$

- d) Prevalensi penderita obesitas (X_4)

$$K_1 = 16,69796 \text{ dan } K_2 = 17,83061$$

- e) Proporsi konsumsi makanan olahan (X_5)

$$K_1 = 30,78163 \text{ dan } K_2 = 32,87551$$

Diperoleh pemodelan regresi nonparametrik *spline truncated* dengan menggunakan dua titik knot sebagai berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - 24,89796)_+ + \hat{\beta}_3 (x_1 - 26,93061)_+ + \hat{\beta}_4 x_2 + \hat{\beta}_5 (x_2 - 34,08163)_+ + \hat{\beta}_6 (x_2 - 36,17551)_+ + \hat{\beta}_7 x_3 + \hat{\beta}_8 (x_3 - 2,165306)_+ + \hat{\beta}_9 (x_3 - 2,820408)_+ + \hat{\beta}_{10} x_4 + \hat{\beta}_{11} (x_4 - 16,69796)_+ + \hat{\beta}_{12} (x_4 - 17,83061)_+ + \hat{\beta}_{13} x_5 + \hat{\beta}_{14} (x_5 - 30,78163)_+ + \hat{\beta}_{15} (x_5 - 32,87551)_+$$

Estimasi paramater model regresi nonparametrik *spline truncated* dengan dua titik knot dapat dilihat pada Tabel 5 di bawah.

Tabel 5. Nilai GCV dengan Tiga Titik Knot

X_1	X_2	X_3	X_4	X_5	GCV
24,897	34,081	2,165	16,697	30,781	
26,930	36,175	2,820	17,830	32,875	53,148
33,706	43,155	5,004	21,606	39,855	
24,897	34,081	2,165	16,697	30,781	
26,930	36,175	2,820	17,830	32,875	53,184
33,028	42,457	4,785	21,228	39,157	
24,897	34,081	2,165	16,697	30,781	
27,608	36,175	3,038	18,208	33,573	53,835
33,706	43,155	5,004	21,606	39,855	
24,897	34,081	2,165	16,697	30,781	
27,608	36,873	3,038	18,208	33,573	54,010
33,028	42,457	4,785	21,228	39,157	
24,897	34,081	2,165	16,697	30,781	
26,930	36,175	2,820	17,830	32,875	54,072
32,351	41,759	4,567	20,851	38,459	

Berdasarkan Tabel 5. didapatkan nilai GCV paling minimum unuk model regresi nonparametrik *spline truncated* dengan tiga titik knot adalah 53,148. Titik knot optimal untuk masing-masing variabel adalah sebagai berikut.

- a) Proporsi konsumsi makanan manis (X_1)

$$K_1 = 24,897, K_2 = 26,930, \text{ dan } K_3 = 33,706$$

- b) Proporsi aktivitas fisik (X_2)

$$K_1 = 34,081, K_2 = 36,175, \text{ dan } K_3 = 43,155$$

- c) Proporsi konsumsi rokok tembakau (X_3)

$$K_1 = 2,165, K_2 = 2,820, \text{ dan } K_3 = 5,004$$

- d) Prevalensi penderita obesitas (X_4)

$$K_1 = 16,697, K_2 = 17,830, \text{ dan } K_3 = 21,6061$$

- e) Proporsi konsumsi makanan olahan (X_5)

$K_1=30,7816$, $K_2 = 32,8755$, dan $K_3 = 39,8551$
 Diperoleh pemodelan regresi nonparametrik *spline truncated* dengan menggunakan tiga titik knot sebagai berikut.

$$\hat{y} = -29,235 + 4,890x_1 - 12,770(x_1 - 24,8979)_+ + 7,547(x_1 - 26,9306)_+ + 0,987(x_1 - 33,7061) + 2,343x_2 + 0,221(x_2 - 34,0816)_+ - 3,768(x_2 - 36,1755)_+ + 0,305(x_2 - 43,1551) + 13,952x_3 - 96,424(x_3 - 2,1653)_+ + 92,643(x_3 - 2,8204)_+ - 12,430(x_3 - 5,0040) - 4,156x_4 - 13,977(x_4 - 16,6979)_+ + 21,822(x_4 - 17,8306)_+ - 2,904(x_4 - 21,6061) - 1,467x_5 - 6,083(x_5 - 30,7816)_+ + 11,730(x_5 - 32,8755)_+ - 4,204(x_5 - 39,8551)_+$$

PEMILIHAN MODEL TERBAIK

Pada pemodelan regresi nonparametrik *spline truncated* akan dicobakan beberapa titik knot yaitu satu titik knot, dua titik knot, dan tiga titik knot. Berikut ini ditunjukkan titik knot beserta nilai GCV untuk pemilihan model terbaik pada Tabel 6.

Tabel 6. Hasil Pemilihan Titik Knot Optimal

Banyak titik knot	Knot Optimum					GCV
	X_1	X_2	X_3	X_4	X_5	
1	22,865	31,988	1,510	15,565	28,688	89,331
2	24,898	34,082	2,165	16,698	30,782	61,658
	26,931	36,176	2,820	17,831	32,876	
3	24,898	34,082	2,165	16,698	30,782	53,148
	26,931	36,176	2,820	17,831	32,876	
	33,706	43,155	5,004	21,606	39,855	

Berdasarkan Tabel 6. diperoleh bahwa nilai GCV yang paling minimum terdapat pada model *spline* dengan menggunakan tiga titik knot yaitu sebesar 53,148, sehingga pemodelan diabetes melitus tipe 2 dan beberapa variabel yang mempengaruhinya akan menggunakan regresi nonparametrik *spline truncated* dengan tiga titik knot.

PENGUJIAN PARAMETER MODEL REGRESI

Pengujian parameter secara serentak dilakukan dengan taraf signifikansi sebesar 5% dengan menggunakan hipotesis statistik uji sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{20} = 0$$

$$H_1 : \text{minimal ada satu } \beta_j \neq 0, j = 1, 2, \dots, 20$$

Regresi	20	4932,873	246,643	9,395	0,000
Error	17	446,281	26,251		
Total	37	5379,154			

Berdasarkan Tabel 7, didapatkan nilai statistik uji $F_{hitung} = 9,395$ dengan nilai $p - value = 0,000$. Nilai $F_{hitung} = 9,395 \geq F_{(0,05;20;17)} = 2,23$ dan $p - value = 0,000 < \alpha(0,05)$, sehingga keputusan yang diperoleh adalah Tolak H_0 . Hal ini menunjukkan bahwa terdapat minimal satu variabel independen yang berpengaruh signifikan terhadap variabel dependen. Oleh karena itu dapat dilanjutkan untuk pengujian parameter secara parsial. Hasil pengujian parameter secara parsial ditampilkan pada Tabel 8.

Tabel 8. Hasil Pengujian Parameter Secara Parsial

Variabel	Parameter Estimator	T	P-value	Keputusan		
Konstan	β_0	-29,235	-0,314	0,757	Tidak Signifikan	
	β_1	4,890	3,302	0,004	Signifikan	
	β_2	-12,770	-2,725	0,014	Signifikan	
	X_1	β_3	7,547	1,846	0,082	Tidak Signifikan
		β_4	0,987	1,076	0,296	Tidak Signifikan
X_2	β_5	2,343	2,019	0,059	Tidak Signifikan	
	β_6	0,221	0,052	0,958	Tidak Signifikan	
	β_7	-3,768	-0,903	0,378	Tidak Signifikan	
	β_8	0,305	0,289	0,775	Tidak Signifikan	
X_3	β_9	13,952	4,533	0,000	Signifikan	
	β_{10}	-96,424	-3,701	0,001	Signifikan	
	β_{11}	92,643	3,117	0,006	Signifikan	
	β_{12}	-12,430	-1,579	0,132	Tidak Signifikan	
X_4	β_{13}	-4,156	-1,369	0,188	Tidak Signifikan	
	β_{14}	-13,977	-1,587	0,130	Tidak Signifikan	
	β_{15}	21,822	2,980	0,008	Signifikan	
	β_{16}	-2,904	-1,975	0,064	Tidak Signifikan	
X_5	β_{17}	-1,467	-1,056	0,305	Tidak Signifikan	
	β_{18}	-6,083	-1,145	0,267	Tidak Signifikan	
	β_{19}	11,730	2,398	0,028	Signifikan	
	β_{20}	-4,204	-3,616	0,002	Signifikan	

Tabel 7. ANOVA Model Regresi *Spline Truncated*

Sumber Variasi	df	SS	MS	F-hit	P-Value
----------------	----	----	----	-------	---------

Tabel 8. merupakan hasil pengujian parameter model secara parsial. Jika nilai $|t_{hitung}| \geq t_{(0,025;17)}$ (2,1098) dan $p - value < \alpha(0,05)$ maka keputusannya adalah tolak H_0 . Berdasarkan Tabel 8. diketahui bahwa terdapat 8 parameter keputusannya adalah tolak H_0 yang berarti bahwa parameter tersebut signifikan terhadap model, sedangkan 12 parameter lainnya tidak signifikan terhadap model. Namun, meski terdapat parameter yang tidak signifikan, variabel tersebut masih digunakan karena minimal terdapat satu parameter yang signifikan, sehingga variabel independen X_1, X_3, X_4 , dan X_5 memberikan pengaruh yang signifikan terhadap persentase diabetes melitus tipe 2 di Indonesia.

INTERPRETASI MODEL REGRESI NONPARAMETRIK SPLINE TRUNCATED

Berikut merupakan interpretasi dari model nonparametrik *spline truncated* pada setiap variabel yang berpengaruh signifikan terhadap persentase diabetes melitus tipe 2 di Indonesia.

1) Model *Spline* pada Variabel Proporsi Konsumsi Makanan Manis (X_1)

$$\hat{y} = -29,235 + 4,890x_1 - 12,770(x_1 - 24,8979)_+ + 7,547(x_1 - 26,9306)_+ + 0,987(x_1 - 33,7061)_+ + \begin{cases} 4,890x_1, & x_1 < 24,8979 \\ 317,946 - 7,880x_1, & 24,8979 \leq x_1 < 26,9306 \\ 114,701 - 0,333x_1, & 26,9306 \leq x_1 < 33,7061 \\ 81,433 + 0,654x_1, & x_1 \geq 33,7061 \end{cases}$$

Berdasarkan model di atas, dapat diketahui bahwa pada variabel proporsi konsumsi makanan manis (X_1) terdapat empat interval, dimana untuk setiap interval memiliki interpretasi model yang berbeda. Interval pertama merupakan proporsi konsumsi makanan manis yang kurang dari 24,8979 persen, yang berarti bahwa apabila proporsi konsumsi makanan manis naik satu persen maka persentase diabetes melitus tipe 2 meningkat sebesar 4,890 persen. Interval kedua merupakan proporsi konsumsi makanan manis yang lebih dari sama dengan 24,8979 persen hingga kurang dari 26,9306 persen, yang berarti bahwa apabila proporsi konsumsi makanan manis naik satu persen maka persentase diabetes melitus tipe 2 turun sebesar 7,880 persen. Interval ketiga merupakan proporsi konsumsi makanan manis yang lebih dari sama

dengan 26,9306 persen hingga kurang dari 33,7061 persen, yang berarti bahwa apabila proporsi konsumsi makanan manis naik satu persen maka persentase diabetes melitus tipe 2 turun sebesar 0,333 persen. Interval keempat merupakan proporsi konsumsi makanan manis yang lebih dari sama dengan 33,7061 persen, yang berarti bahwa apabila proporsi konsumsi makanan manis naik satu persen maka persentase diabetes melitus tipe 2 meningkat sebesar 0,654 persen.

2) Model *Spline* pada Variabel Proporsi Konsumsi Rokok Tembakau (X_3)

$$\hat{y} = -29,235 + 13,952x_3 - 96,424(x_3 - 2,1653)_+ + 92,643(x_3 - 2,8204)_+ - 12,430(x_3 - 5,0040)_+ + \begin{cases} 13,952x_3 & ; & x_3 < 2,1653 \\ 208,787 - 82,472x_3 & ; & 2,1653 \leq x_3 < 2,8204 \\ -52,503 + 10,171x_3 & ; & 2,8204 \leq x_3 < 5,0040 \\ 9,696 - 2,259x_3 & ; & x_3 \geq 5,0040 \end{cases}$$

Berdasarkan model di atas, dapat diketahui bahwa pada variabel proporsi konsumsi rokok tembakau (X_3) terdapat empat interval, dimana untuk setiap interval memiliki interpretasi model yang berbeda. Interval pertama merupakan proporsi konsumsi rokok tembakau yang kurang dari 2,1653 persen, yang berarti bahwa apabila proporsi konsumsi rokok tembakau naik satu persen maka persentase diabetes melitus tipe 2 meningkat sebesar 13,952 persen. Interval kedua merupakan proporsi konsumsi rokok tembakau yang lebih dari sama dengan 2,1653 persen hingga kurang dari 2,8204 persen, yang berarti bahwa apabila proporsi konsumsi rokok tembakau naik satu persen maka persentase diabetes melitus tipe 2 turun sebesar 82,472 persen. Interval ketiga merupakan proporsi konsumsi rokok tembakau yang lebih dari sama dengan 2,8204 persen hingga kurang dari 5,0040 persen, yang berarti bahwa apabila proporsi konsumsi rokok tembakau naik satu persen maka persentase diabetes melitus tipe 2 meningkat sebesar 10,171 persen. Interval keempat merupakan proporsi konsumsi rokok tembakau yang lebih dari sama dengan 5,0040 persen, yang berarti bahwa apabila proporsi konsumsi rokok tembakau naik satu persen maka persentase diabetes melitus tipe 2 turun sebesar 2,259 persen.

3) Model *Spline* pada Variabel Prevalensi Penderita Obesitas (X_4)

$$\hat{y} = -29,235 - 4,156x_4 - 13,977(x_4 - 16,6979)_+ + 21,822(x_4 - 17,8306)_+ - 2,904(x_4 - 21,6061)_+ \\ = \begin{cases} -4,156x_4 & ; & x_4 < 16,6979 \\ 233,387 - 18,133x_4 & ; & 16,6979 \leq x_4 < 17,8306 \\ -155,713 + 3,689x_4 & ; & 17,8306 \leq x_4 < 21,6061 \\ -92,968 + 0,785x_4 & ; & x_4 \geq 21,6061 \end{cases}$$

Berdasarkan model di atas, dapat diketahui bahwa pada variabel prevalensi penderita obesitas (X_4) terdapat empat interval, dimana untuk setiap interval memiliki interpretasi model yang berbeda. Interval pertama merupakan prevalensi penderita obesitas yang kurang dari 16,6979 persen, yang berarti bahwa apabila prevalensi penderita obesitas naik satu persen maka persentase diabetes melitus tipe 2 turun sebesar 4,156 persen. Interval kedua merupakan prevalensi penderita obesitas yang lebih dari sama dengan 16,6979 persen hingga kurang dari 17,8306 persen, yang berarti bahwa apabila prevalensi penderita obesitas naik satu persen maka persentase diabetes melitus tipe 2 turun sebesar 18,133 persen. Interval ketiga merupakan prevalensi penderita obesitas yang lebih dari sama dengan 17,8306 persen hingga kurang dari 21,6061 persen, yang berarti bahwa apabila prevalensi penderita obesitas naik satu persen maka persentase diabetes melitus tipe 2 meningkat sebesar 3,689 persen. Interval keempat merupakan prevalensi penderita obesitas yang lebih dari sama dengan 21,6061 persen, yang berarti bahwa apabila prevalensi penderita obesitas naik satu persen maka persentase diabetes melitus tipe 2 meningkat sebesar 0,785 persen.

4) Model *Spline* pada Variabel Proporsi Konsumsi Makanan Olahan (X_5)

$$\hat{y} = -29,235 - 1,467x_5 - 6,083(x_5 - 30,7816)_+ + 11,730(x_5 - 32,8755)_+ - 4,204(x_5 - 39,8551)_+ \\ = \begin{cases} -1,467x_5 & ; & x_5 < 30,7816 \\ 187,244 - 7,550x_5 & ; & 30,7816 \leq x_5 < 32,8755 \\ -198,385 + 4,180x_5 & ; & 32,8755 \leq x_5 < 39,8551 \\ -30,834 - 0,024x_5 & ; & x_5 \geq 39,8551 \end{cases}$$

Berdasarkan model di atas, dapat diketahui bahwa pada variabel proporsi konsumsi makanan olahan (X_5) terdapat empat interval, dimana untuk setiap interval memiliki

interpretasi model yang berbeda. Interval pertama merupakan proporsi konsumsi makanan olahan yang kurang dari 30,7816 persen, yang berarti bahwa apabila proporsi konsumsi makanan olahan naik satu persen maka persentase diabetes melitus tipe 2 turun sebesar 1,467 persen. Interval kedua merupakan proporsi konsumsi makanan olahan yang lebih dari sama dengan 30,7816 persen hingga kurang dari 32,8755 persen, yang berarti bahwa apabila proporsi konsumsi makanan olahan naik satu persen maka persentase diabetes melitus tipe 2 turun sebesar 7,550 persen. Interval ketiga merupakan proporsi konsumsi makanan olahan yang lebih dari sama dengan 32,8755 persen hingga kurang dari 39,8551 persen, yang berarti bahwa apabila proporsi konsumsi makanan olahan naik satu persen maka persentase diabetes melitus tipe 2 meningkat sebesar 4,180 persen. Interval keempat merupakan proporsi konsumsi makanan olahan yang lebih dari sama dengan 39,8551 persen, yang berarti bahwa apabila proporsi konsumsi makanan olahan naik satu persen maka persentase diabetes melitus tipe 2 turun sebesar 0,024 persen.

PENUTUP

SIMPULAN

Model regresi nonparametrik *spline truncated* terbaik untuk memodelkan kasus diabetes melitus tipe 2 di Indonesia adalah dengan tiga titik knot dengan nilai GCV minimum 53,148 dan nilai koefisien determinasi yang dihasilkan model tersebut sebesar 91,55 persen. Dari kelima variabel yang digunakan, terdapat empat variabel yang berpengaruh signifikan yaitu proporsi konsumsi makanan manis (X_1), proporsi konsumsi rokok tembakau (X_3), prevalensi penderita obesitas (X_4), dan proporsi konsumsi makanan olahan (X_5). Berikut merupakan hasil pemodelan yang diperoleh regresi nonparametrik *spline truncated*.

$$\hat{y} = -29,235 + 4,890x_1 - 12,770(x_1 - 24,8979)_+ + 7,547(x_1 - 26,9306)_+ + 0,987(x_1 - 33,7061)_+ + 2,343x_2 + 0,221(x_2 - 34,0816)_+ - 3,768(x_2 - 36,1755)_+ + 0,305(x_2 - 43,1551) + 13,952x_3 - 96,424(x_3 - 2,1653)_+ + 92,643(x_3 - 2,8204)_+ - 12,430(x_3 - 5,0040) - 4,156x_4 - 13,977(x_4 - 16,6979)_+ + 21,822(x_4 - 17,8306)_+ - 2,904(x_4 - 21,6061) - 1,467x_5 - 6,083(x_5 - 30,7816)_+ + 11,730(x_5 - 32,8755)_+ - 4,204(x_5 - 39,8551)_+$$

SARAN

Sebagai upaya untuk menekan angka kejadian diabetes melitus di Indonesia, disarankan agar pemerintah fokus pada penyusunan program pencegahan yang menargetkan faktor-faktor signifikan seperti konsumsi makanan manis, rokok tembakau, obesitas, dan konsumsi makanan olahan. Program tersebut dapat berupa kampanye nasional tentang bahaya konsumsi gula berlebih dan rokok, pemberian insentif untuk produk makanan sehat, serta peningkatan akses masyarakat terhadap fasilitas olahraga dan layanan konsultasi gizi. Penggunaan metode regresi nonparametrik *spline truncated* dalam penelitian ini memberikan keunggulan dalam memodelkan hubungan yang tidak linier antar variabel-variabel independen dan persentase diabetes melitus tipe 2. Sebagai pengembangan ke depannya, penelitian serupa dapat menggunakan pendekatan regresi nonparametrik lain seperti Regresi Kernel, *K-nearest Neighbors Regression* (KNN regression), Regresi B-Spline, dan Regresi P-Spline untuk dibandingkan dalam mengevaluasi tingkat akurasi model.

DAFTAR PUSTAKA

- Alamsyah, I. F., Esra, R., Awalia, S., & Nohe, D. A. (2022). ANALISIS REGRESI DATA PANEL UNTUK MENGETAHUI FAKTOR YANG MEMENGARUHI JUMLAH PENDUDUK MISKIN DI KALIMANTAN TIMUR. *Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya Terbitan II, Mei 2022, Samarinda, Indonesia*.
- Association, A. D. (2020). Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2020. *PubMed*.
- Budiantara, I. N. (2011). Penelitian Bidang Regresi Spline Menuju Terwujudnya Penelitian Statistika yang Mandiri dan Berkarakter. *Seminar Nasional FMIPA Undiksha*.
- Budiantara, I. N. (2019). *Regresi Nonparametrik Spline Truncated*. Surabaya: ITS Press.
- Delfina, S., M., I., C. H., & Ayatillahi, S. (2021). Analisis Determinan Faktor Risiko Kejadian Diabetes Mellitus Tipe 2 Pada Usia Produktif. *JURNAL KESEHATAN TAMBUSAI*.
- Draper, N. R., & Smith, H. (1998). *Applied Regression and Spline Smoothing (3rd Edition)*. New York.
- Eubank, R. L. (1999). *Nonparametric Regression and Spline*. New York: Marcel Dekker Inc.
- Harrel, F. E. (2015). *Regression Modeling Strategies: with Application to Linear Models, Logistic and Ordinal Regression, and Survival Analysis (2nd ed.)*. Springer.
- Kurniawati, N. A., & Budiantara, I. N. (2019). Pemodelan Tingkat Pengangguran Terbuka di Provinsi Jawa Barat Menggunakan Regresi Nonparametrik Spline truncated. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*.
- Miranda, T. G., & Sari, L. Y. (2023). FAKTOR RISIKO KEJADIAN DIABETES MELITUS (DM) TIPE 2 DI POLIKLINIK PENYAKIT DALAM RS KOTA BENGKULU. FAKTOR RISIKO KEJADIAN DIABETES MELITUS (DM) TIPE 2 DI POLIKLINIK PENYAKIT DALAM RS KOTA BENGKULU.
- PERKENI. (2021). *PENGELOLAAN DAN PENCEGAHAN DIABETES MELITUS TIPE 2 DEWASA DI INDONESIA 2021*. PB. PERKENI.
- Putri, I. M. (2018). REGRESI NONPARAMETRIK SPLINE TRUNCATED UNTUK MEMODELKAN PERSENTASE PESERTA KB BARU PASCA PERSALINAN DAN PASCA KEGUGURAN DI JAWA TIMUR TAHUN 2016. .
- Ramadhan, R., Irmeilyana, & Desiani, A. (2024). PENERAPAN REGRESI NONPARAMETRIK SPLINE TRUNCATED DALAM MEMODELKAN HUBUNGAN FAKTOR-FAKTOR YANG mempengaruhi INDEKS PEMBANGUNAN GENDER (IPG) DI PROVINSI SUMATERA UTARA. *JURNAL SAINS TERAPAN*.
- Safire, E. D., & Purhadi. (2023). Pemodelan Faktor-Faktor yang mempengaruhi Jumlah Kasus Diabetes Melitus di Jawa Timur Menggunakan GWGPR dan GWNBR.
- Sun, H., Pouya, S., Suvi, K., Moritz, P., Katherine, O., B., D., & Abdul, B. (2021). IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *HHS Public Access*.
- Wongkar, D. C., Ruliana, S., & M., F. (2023). Analisis Regresi Nonparametrik Spline Truncated untuk Menganalisis Faktor-Faktor yang mempengaruhi Tingkat Pengangguran Terbuka di Provinsi Sulawesi Selatan. *Journal of Statistics and Its Application on Teaching and Research*.