

BAB I PENDAHULUAN

1.1 Latar Belakang

Inflasi merupakan salah satu indikator makroekonomi yang selalu menjadi perhatian utama pemerintah dan bank sentral karena berpengaruh pada stabilitas harga, daya beli masyarakat, dan arah kebijakan moneter (Triwahyuni, 2021). Di Indonesia, inflasi menunjukkan dinamika yang cukup beragam dalam beberapa tahun terakhir. Berdasarkan data Badan Pusat Statistika (BPS) dan Bank Indonesia (BI), laju inflasi tahunan (*year-on-year*) Indonesia sempat menurun tajam dari 3,13% pada 2020 menjadi 1,56% pada 2021, kemudian kembali meningkat menjadi 5,51% pada 2022, dan melandai ke 2,61% pada akhir 2023. Hingga 2025, inflasi tercatat sebesar 2,31%, dengan deflasi yang terjadi pada Februari 2025 sebesar -0,09% (Bank Indonesia, 2025). Berdasarkan data tersebut, dapat disimpulkan bahwa inflasi di Indonesia tidak hanya fluktuatif dalam jangka panjang, tetapi juga memiliki pola nonlinear yang kompleks.

Kompleksitas data inflasi tidak terlepas dari banyaknya faktor ekonomi yang memengaruhinya. Menurut Ningrum dkk. (2025), kondisi ini menyebabkan tingginya kemungkinan terjadi multikolinearitas dalam pemodelan inflasi, sehingga analisis dengan metode klasik menjadi kurang memadai. Oleh karena itu, diperlukan metode yang mampu menyederhanakan data tanpa menghilangkan informasi utama. Salah satu metode yang dapat digunakan adalah *Principal Component Analysis* (PCA), yang mereduksi variabel-variabel berkorelasi menjadi sejumlah komponen utama yang saling bebas (Jolliffe & Cadima, 2016).

Selain multikolinearitas, pola hubungan antarvariabel juga tidak teratur dan tidak mengikuti pola parametrik, sehingga sulit dianalisis hanya dengan menggunakan metode PCA ataupun regresi parametrik berbasis PCA (Islamiyati dkk., 2022). Oleh karena itu, diperlukan metode lain yang lebih fleksibel, yaitu regresi nonparametrik *B-spline*. Metode tersebut tidak bergantung pada asumsi bentuk hubungan tertentu, sehingga mampu menangkap pola nonlinear yang muncul pada data inflasi. Selain itu, *B-spline* lebih stabil dibandingkan *spline* standar, karena dapat mengurangi masalah hampir singular pada matriks akibat pemilihan titik knot yang kurang tepat (Raupong, 2009).

Beberapa penelitian terdahulu telah membahas pemodelan inflasi menggunakan berbagai pendekatan statistik. Misalnya, Melius dan Wijaya (2023) menggunakan regresi linear untuk menganalisis faktor-faktor yang memengaruhi inflasi di Indonesia dan menemukan bahwa beberapa variabel tidak signifikan atau berpengaruh berbeda dari yang diprediksi teori. Selanjutnya, Ningrum dkk. (2025) menerapkan regresi komponen utama untuk mengatasi multikolinearitas, namun model yang diperoleh belum mampu menjelaskan inflasi Indonesia pada periode 2020-2022 secara optimal. Sementara itu, metode regresi nonparametrik menunjukkan hasil yang lebih baik. Misalnya, Andrianzah (2023) menggunakan regresi nonparametrik *spline truncated* untuk memodelkan inflasi Indonesia dan memperoleh nilai koefisien determinasi sebesar 92,80%, serta Suwarno (2023) menggunakan regresi nonparametrik *B-spline* dan memperoleh nilai koefisien determinasi sebesar 87,53%. Selain itu, Rahasia dkk. (2020) juga menunjukkan bahwa metode *B-spline* mampu memberikan pemodelan yang fleksibel dan akurat pada data nilai tukar mata uang rupiah terhadap dolar Amerika.

Akan tetapi, penelitian-penelitian terdahulu tersebut umumnya masih menggunakan metode secara terpisah. Penelitian dengan regresi linear dan regresi nonparametrik belum mempertimbangkan secara khusus adanya multikolinearitas pada variabel prediktor, sedangkan penelitian yang menggunakan PCA belum mampu menangani hubungan nonlinear pada data inflasi. Berdasarkan hal tersebut, penelitian ini menggunakan model regresi nonparametrik *B-spline* dengan PCA pada data inflasi Indonesia. Penggabungan kedua metode ini diharapkan dapat memberikan hasil pemodelan yang lebih sesuai dengan karakteristik data inflasi yang kompleks.

1.2 Batasan Masalah

Batasan masalah digunakan untuk menghindari adanya penyimpangan dan pelebaran pokok masalah agar memudahkan dalam pembahasan sehingga tujuan penelitian akan tercapai. Berikut beberapa batasan masalah dalam penelitian ini:

1. Data yang digunakan merupakan data sekunder mengenai inflasi dan variabel-variabel prediktor yang relevan, dengan rentang waktu Januari 2015 hingga Desember 2024 dalam bentuk data perbulan.
2. Pemilihan jumlah komponen utama pada PCA didasarkan pada persentase varians kumulatif $\geq 75\%$.
3. Titik knot optimal yang digunakan dalam model *B-spline* dipilih berdasarkan nilai *Generalized Cross Validation* (GCV) minimum.

1.3 Tujuan Penelitian

Sejalan dengan latar belakang yang telah dituliskan, maka tujuan yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Memperoleh estimasi parameter model regresi nonparametrik *B-spline* dengan PCA.
2. Memperoleh komponen utama sebagai variabel prediktor baru melalui reduksi dimensi menggunakan PCA.
3. Memperoleh model data inflasi Indonesia menggunakan regresi nonparametrik *B-spline* dengan PCA.

1.4 Manfaat Penelitian

Berikut adalah beberapa manfaat yang dapat diperoleh dari penelitian ini:

1. Menambah pengetahuan dan wawasan pada pengembangan metode analisis data ekonomi, khususnya pengimplementasian regresi nonparametrik *B-spline* dengan PCA pada data inflasi Indonesia.
2. Memberikan gambaran secara statistika mengenai hubungan inflasi dan variabel-variabel prediktornya, sehingga dapat menjadi referensi bagi mahasiswa atau peneliti selanjutnya yang tertarik pada kajian inflasi.

1.5 Landasan Teori

1.5.1 Matriks Korelasi

Matriks adalah suatu susunan bilangan atau fungsi yang disajikan dalam baris dan kolom dan diapit oleh dua tanda kurung. Bilangan atau fungsi tersebut dinamakan entri atau elemen dalam matriks (Marsudi & Marjono, 2012). Ukuran suatu matriks atau ordo dijelaskan dengan menuliskan banyak baris (n) dan banyak kolom (p). Misalkan X suatu matriks berukuran $n \times p$, maka matriks tersebut secara umum dapat dituliskan sebagai berikut:

$$X_{n \times p} = [x_{ij}] = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

di mana x_{ij} menyatakan elemen baris i dengan $i = 1, 2, \dots, n$ dan kolom j dengan $j = 1, 2, \dots, p$.

Matriks korelasi (R) adalah matriks simetri berukuran $p \times p$ yang diperoleh dengan menormalisasi kovarians terhadap standar deviasi masing-masing variabel. Matriks korelasi (R) diperoleh dengan menggunakan Persamaan (1) berikut (Tahir, 2021):

$$\rho_{hj} = \frac{\sigma_{hj}}{\sqrt{\sigma_{hh} \cdot \sigma_{jj}}}; h = 1, 2, \dots, p; j = 1, 2, \dots, p \quad (1)$$

dengan ρ_{hj} menunjukkan koefisien korelasi yang dinyatakan dalam varians (σ_{hh} dan σ_{jj}) dan kovarians (σ_{hj}).

dan dapat dinyatakan dalam bentuk matriks pada Persamaan (2) berikut:

$$R = \begin{bmatrix} \frac{\sigma_{11}}{\sqrt{\sigma_{11} \cdot \sigma_{11}}} & \frac{\sigma_{12}}{\sqrt{\sigma_{11} \cdot \sigma_{22}}} & \cdots & \frac{\sigma_{1p}}{\sqrt{\sigma_{11} \cdot \sigma_{pp}}} \\ \frac{\sigma_{21}}{\sqrt{\sigma_{22} \cdot \sigma_{11}}} & \frac{\sigma_{22}}{\sqrt{\sigma_{22} \cdot \sigma_{22}}} & \cdots & \frac{\sigma_{2p}}{\sqrt{\sigma_{22} \cdot \sigma_{pp}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{p1}}{\sqrt{\sigma_{pp} \cdot \sigma_{11}}} & \frac{\sigma_{p2}}{\sqrt{\sigma_{pp} \cdot \sigma_{22}}} & \vdots & \frac{\sigma_{pp}}{\sqrt{\sigma_{pp} \cdot \sigma_{pp}}} \end{bmatrix} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \vdots & 1 \end{bmatrix} \quad (2)$$

Setiap elemen diagonal utama bernilai 1 karena korelasi antara suatu variabel dengan dirinya sendiri selalu sempurna.

1.5.2 Nilai Eigen dan Vektor Eigen

Jika A adalah matriks bujur sangkar berukuran $n \times n$, maka ada nilai eigen (λ) dan vektor eigen (e) tak nol di dalam R^n yang bersesuaian dengan λ sehingga memenuhi Persamaan (3) berikut (Pranata & Gunawan, 2018):

$$Ae = \lambda e \quad (3)$$

Kemudian dapat dituliskan menjadi Persamaan (4) berikut:

$$Ae = \lambda Ie$$

$$\begin{aligned} A\mathbf{e} - \lambda I\mathbf{e} &= 0 \\ (A - \lambda I)\mathbf{e} &= 0 \end{aligned} \quad (4)$$

dengan I adalah matriks identitas yang berukuran sama dengan matriks A , dalam matriks dapat dituliskan:

$$A_{n \times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \vdots & a_{nn} \end{bmatrix}, I_{n \times n} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & 1 \end{bmatrix}, \mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

Karena $\mathbf{e} \neq 0$, maka:

$$\det(A - \lambda I) = 0 \quad (5)$$

Nilai skalar yang memenuhi Persamaan (5) disebut nilai-nilai eigen dari matriks A . Sedangkan vektor eigen yang bersesuaian dengan suatu nilai eigen dapat ditemukan dengan mensubstitusikan nilai eigen tersebut ke dalam Persamaan (4).

1.5.3 Principal Component Analysis (PCA)

Salah satu asumsi klasik dari regresi linear adalah tidak terjadi multikolinearitas antarvariabel prediktor. Multikolinearitas terjadi ketika terdapat hubungan linear yang kuat antarvariabel prediktor, sehingga dapat menyebabkan estimasi parameter menjadi tidak stabil, varians dari koefisien regresi meningkat, serta sulit untuk menginterpretasikan pengaruh masing-masing variabel prediktor secara individual (Amni dkk., 2024). Deteksi multikolinearitas dapat dilakukan melalui analisis matriks korelasi antarvariabel prediktor maupun menggunakan *Variance Inflation Factor* (VIF). Adapun kriteria multikolinearitas berdasarkan nilai VIF adalah sebagai berikut:

- Apabila nilai $VIF < 5$ atau $VIF \approx 1$, maka multikolinearitas tergolong lemah (Octaviana dkk., 2019).
- Apabila nilai $VIF \geq 5$, maka multikolinearitas tergolong sedang (Utami dkk., 2013).
- Apabila nilai $VIF \geq 10$, maka multikolinearitas tergolong kuat (Maati dkk., 2025).

Pada dasarnya tidak terdapat ambang batas universal untuk nilai VIF, beberapa penelitian menekankan bahwa kriteria tersebut tidak sebaiknya diterapkan secara kaku. Oleh karena itu, kriteria VIF yang digunakan untuk mendeteksi multikolinearitas antar variabel prediktor dapat bervariasi antar peneliti (Winter, 2025). Nilai VIF dapat dihitung menggunakan rumus pada Persamaan (6) berikut (Sartika dkk., 2020):

$$VIF_j = \frac{1}{1 - R_j^2}; j = 1, 2, \dots, p \quad (6)$$

dengan R_j^2 merupakan koefisien determinasi yang diperoleh dari hasil regresi variabel prediktor X_j terhadap variabel prediktor lainnya.

Permasalahan multikolinearitas tersebut dapat diatasi dengan menggunakan metode *Principal Component Analysis* (PCA). *Principal Component Analysis* merupakan salah satu metode analisis data multivariat yang bertujuan untuk menjelaskan struktur varians-kovarians dari sekumpulan variabel melalui pembentukan variabel baru yang disebut komponen utama. Komponen utama tersebut merupakan kombinasi linear dari variabel asal dan bersifat saling bebas (Sriningsih dkk., 2018). Metode PCA digunakan

untuk mereduksi dimensi data atau variabel dengan jumlah besar menjadi lebih sedikit tanpa menghilangkan informasi penting dalam data.

Secara matematis, komponen utama dibentuk dari variabel acak $X_1, X_2, \dots, X_p$ yang memiliki skala pengukuran yang sama. Dari p variabel asal dapat diturunkan sebanyak p komponen utama yang mampu menerangkan seluruh varians total data. Namun, seringkali varians total dapat dijelaskan secara memadai hanya oleh sejumlah kecil komponen utama. Misalkan dipilih sebanyak q komponen utama, dengan $q \leq p$, yang dianggap telah cukup mewakili informasi utama dalam data. Proses pembentukan komponen utama ini menghasilkan pasangan nilai eigen dan vektor eigen yang saling ortonormal, yaitu $(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p)$ dengan $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$. Maka, komponen utama ke- r didefinisikan pada Persamaan (7) sebagai berikut (Amni dkk., 2024):

$$\begin{aligned} W_r &= e_r^T X \\ W_r &= e_{r1}X_1 + e_{r2}X_2 + \dots + e_{rp}X_p; r = 1, 2, \dots, q \\ e_r^T X &= [e_{r1} \quad e_{r2} \quad \dots \quad e_{rp}] \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} \end{aligned} \quad (7)$$

bentuk matriks pada Persamaan (7) adalah sebagai berikut:

$$e_r^T X = \begin{bmatrix} e_{11} & e_{12} & \dots & e_{1p} \\ e_{21} & e_{22} & \dots & e_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ e_{r1} & e_{r2} & \dots & e_{rp} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}$$

dengan W_1 adalah komponen pertama yang diperoleh dengan memaksimumkan nilai $e_1^T \Sigma e_1 = \lambda_1$, W_2 adalah komponen kedua yang dibentuk untuk menjelaskan varians yang tersisa dari komponen pertama dengan memaksimumkan nilai $e_2^T \Sigma e_2 = \lambda_2$, dan W_p adalah komponen ke- p yang menjelaskan sisa varians selain dari $W_1, W_2, \dots, W_{p-1}$ dengan memaksimumkan nilai $e_p^T \Sigma e_p = \lambda_p$.

Pada permasalahan regresi, variabel-variabel yang digunakan sering kali memiliki skala pengukuran yang berbeda, sehingga dapat memengaruhi hasil analisis. Oleh karena itu, PCA didasarkan pada matriks korelasi dan menggunakan variabel yang terstandarisasi untuk mengatasi perbedaan skala tersebut. Standarisasi dari variabel X menjadi variabel X^* dilakukan melalui transformasi Persamaan (8) berikut (Widiharih, 2001):

$$X_j^* = \frac{X_j - \mu_j}{\sigma_j}; j = 1, 2, \dots, p \quad (8)$$

Komponen utama kemudian dibentuk sebagai kombinasi linear dari variabel yang telah distandarisasi, yang secara matematis dinyatakan sebagai Persamaan (9) berikut:

$$\begin{aligned} W_r &= e_r^T X^* \\ W_r &= e_{r1}X_1^* + e_{r2}X_2^* + \dots + e_{rp}X_p^* \end{aligned} \quad (9)$$

Terdapat beberapa kriteria yang dapat digunakan dalam menentukan jumlah komponen utama, salah satunya adalah berdasarkan persentase varians kumulatif terhadap total varians. Persentase varians yang dijelaskan oleh komponen utama ke- j dapat diperoleh melalui Persamaan (10) berikut (Siburian dkk., 2019):

$$\frac{\lambda_j}{\sum_{j=1}^p \lambda_j} \times 100\% = \frac{\lambda_j}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \times 100\% ; j = 1, 2, \dots, p \quad (10)$$

Pemilihan komponen utama biasanya dilakukan dengan melihat persentase varians kumulatif, di mana komponen utama dipertahankan hingga mampu menjelaskan sekitar 70% sampai 80% dari total varians data (Putri & Imro'ah, 2021).

1.5.4 Regresi Nonparametrik *B-Spline*

Analisis regresi merupakan metode statistik yang digunakan untuk mengetahui hubungan antara variabel respon dan variabel prediktor (Prahutama, 2013). Secara umum, model regresi linear dapat dinyatakan sebagai Persamaan (11) berikut:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i + \varepsilon_i \quad (11)$$

dengan y_i merupakan variabel respon, x_i merupakan variabel prediktor, β_0 merupakan intersep, β_i merupakan *slope*, dan ε_i merupakan *error*.

Namun, dalam beberapa kasus, hubungan antara variabel respon dan variabel prediktor tidak dapat dijelaskan dengan bentuk fungsi tertentu. Oleh karena itu, regresi nonparametrik digunakan sebagai metode yang tidak mengasumsikan bentuk fungsi regresi secara spesifik, melainkan hanya mengasumsikan bahwa kurva regresi bersifat halus (*smooth*) (Salam dkk., 2022). Secara umum, model regresi nonparametrik dapat dituliskan seperti Persamaan (12) berikut:

$$y_i = f(x_i) + \varepsilon_i \quad (12)$$

dengan $f(x_i)$ merupakan kurva fungsi regresi yang tidak diketahui bentuknya.

Salah satu metode yang sering digunakan dalam regresi nonparametrik adalah regresi *spline*. Regresi *spline* digunakan ketika hubungan antara variabel respon dan variabel prediktor tidak dapat dijelaskan dengan fungsi sederhana. *Spline* merupakan fungsi potongan berupa polinomial dengan orde tertentu yang dihubungkan pada titik-titik tertentu yang disebut knot, sehingga menghasilkan kurva yang halus dan kontinu pada setiap segmen (Sriningsih & Rahmawati, 2018).

Salah satu jenis *spline* yang umum digunakan adalah *B-spline* (Basis *Spline*). *B-spline* mampu membentuk fungsi regresi dengan basis yang lebih stabil secara numerik dibandingkan *spline* polinomial biasa. Selain itu, *B-spline* memiliki sifat lokal, yaitu perubahan data di sekitar suatu knot hanya memengaruhi kurva di sekitar titik tersebut tanpa mengubah keseluruhan bentuk fungsi. Hal tersebut membuat *B-spline* lebih fleksibel dan efisien dalam memodelkan hubungan nonlinear antara variabel respon dan variabel prediktor.

Menurut Ariesta dkk (2021), model regresi nonparametrik apabila didekati menggunakan fungsi *B-spline* dengan orde m dan k titik knot, dapat dituliskan melalui Persamaan (13) berikut:

$$f(x_i) = \sum_{l=1}^{m+k} \beta_l B_{l,m}(x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (13)$$

dengan β_l merupakan parameter regresi *B-spline* ke- l , $B_{l,m}(x_i)$ merupakan basis *B-spline* orde m , dan ε_i merupakan *error* pada pengamatan ke- i .

Fungsi basis *B-spline* pada orde m dan titik knot k dibentuk secara rekursif melalui Persamaan (14) berikut (Handajani dkk., 2023):

$$B_{l,m}(x_i) = \frac{x_i - K_l}{K_{l+m-1} - K_l} B_{l,m-1}(x_i) + \frac{K_{l+m} - x_i}{K_{l+m} - K_{l+1}} B_{l+1,m-1}(x_i) \quad (14)$$

dengan $l = 1, 2, \dots, m + k$ dan $B_{l,0}(x_i) = \begin{cases} 1, & \text{jika } K_l < x_i \leq K_{l+1} \\ 0, & \text{untuk lainnya.} \end{cases}$

Model *B-spline* dibangun dengan memperhatikan tiga aspek penting, yaitu orde *spline* yang digunakan, jumlah knot yang dipilih, dan lokasi knot dalam model (Rahmawati dkk., 2017). Jika jumlah knot terlalu sedikit maka dapat terjadi *underfitting*, sedangkan jika terlalu banyak knot dapat menyebabkan *overfitting*. Oleh karena itu, diperlukan metode seleksi jumlah knot yang optimal (Fajar & Fajariyanto, 2022).

1.5.5 Estimasi Parameter Regresi Nonparametrik *B-Spline*

Pada regresi nonparametrik *B-spline*, salah satu tahap penting yang harus dilakukan adalah mengestimasi parameter yang nantinya membentuk model regresi. Estimasi diperlukan untuk mendapatkan gambaran hubungan yang paling sesuai antara variabel respon dan variabel prediktor melalui kombinasi linear dari fungsi basis *B-spline*. Untuk mengestimasi parameter β pada Persamaan (13) dengan titik knot sebanyak k didefinisikan matriks:

$$\mathbf{B} = B_{l,m}(x_i); l = 1, 2, \dots, m + k$$

atau dapat ditulis sebagai berikut:

$$\mathbf{B} = \begin{bmatrix} B_{1,m}(x_1) & B_{2,m}(x_1) & \cdots & B_{m+k,m}(x_1) \\ B_{1,m}(x_2) & B_{2,m}(x_2) & \cdots & B_{m+k,m}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ B_{1,m}(x_n) & B_{2,m}(x_n) & \cdots & B_{m+k,m}(x_n) \end{bmatrix}$$

$\mathbf{B}$ adalah matriks *B-spline* berukuran $n \times (m + k)$.

Model regresi Persamaan (13) dapat ditulis menjadi Persamaan (15) berikut:

$$f(x_i) = \beta_1 B_{1,m}(x_i) + \beta_2 B_{2,m}(x_i) + \cdots + \beta_{m+k} B_{m+k,m}(x_i) + \varepsilon_i \quad (15)$$

Berdasarkan Persamaan (15) yang merupakan persamaan dengan univariabel *B-spline*, maka persamaan multivariabel *B-spline* dapat ditulis sebagai Persamaan (16) berikut (Mathworks, 2015):

$$f(x_{i1}, x_{i2}, \dots, x_{ip}) = \sum_{l_1=1}^{m_1+k_1} \beta_{1l_1} B_{l_1,m_1}(x_{i1}) + \sum_{l_2=1}^{m_2+k_2} \beta_{2l_2} B_{l_2,m_2}(x_{i2}) + \cdots + \sum_{l_p=1}^{m_p+k_p} \beta_{pl_p} B_{l_p,m_p}(x_{ip}) + \varepsilon_i \quad (16)$$

dan dapat ditulis menjadi Persamaan (17) berikut:

$$y_i = \beta_{11} B_{1,m_1}(x_{i1}) + \cdots + \beta_{1(m_1+k_1)} B_{m_1+k_1,m_1}(x_{i1}) + \cdots + \beta_{p1} B_{1,m_p}(x_{ip}) + \cdots + \beta_{p(m_p+k_p)} B_{m_p+k_p,m_p}(x_{ip}) + \varepsilon_i \quad (17)$$

Model *B-spline* Persamaan (17) jika dibuat dalam bentuk matriks menjadi:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} B_{1,m}(x_{11}) & \cdots & B_{m+k,m}(x_{11}) & \cdots & B_{1,m}(x_{1p}) & \cdots & B_{m+k,m}(x_{1p}) \\ B_{1,m}(x_{21}) & \cdots & B_{m+k,m}(x_{21}) & \cdots & B_{1,m}(x_{2p}) & \cdots & B_{m+k,m}(x_{2p}) \\ \vdots & & \vdots & & \vdots & & \vdots \\ B_{1,m}(x_{n1}) & \cdots & B_{m+k,m}(x_{n1}) & \cdots & B_{1,m}(x_{np}) & \cdots & B_{m+k,m}(x_{np}) \end{bmatrix} \begin{bmatrix} \beta_{11} \\ \beta_{12} \\ \vdots \\ \beta_{p(m+k)} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dan dapat ditulis menjadi Persamaan (18) berikut:

$$\mathbf{y} = \mathbf{B}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (18)$$

Pada penelitian ini, parameter $\boldsymbol{\beta}$ diestimasi menggunakan metode *Ordinary Least Square* (OLS), yaitu dengan meminimumkan jumlah kuadrat *error* atau *Sum of Squared Error* (SSE), yang dirumuskan sebagai berikut (Ariesta dkk., 2021):

$$\begin{aligned} S(\boldsymbol{\beta}) &= \sum_{i=1}^n \varepsilon_i^2 \\ &= \sum_{i=1}^n \left(y_i - \sum_{l=1}^{m+k} \beta_l B_{l,m}(x_i) \right)^2 \end{aligned}$$

Nilai SSE akan diminimumkan dengan cara diturunkan terhadap $\boldsymbol{\beta}$ dan setiap turunannya disamadengankan dengan nol, sehingga solusi estimasi OLS dapat dituliskan secara ringkas dalam bentuk matriks pada Persamaan (19) berikut:

$$\hat{\boldsymbol{\beta}} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{y} \quad (19)$$

dengan $\hat{\boldsymbol{\beta}} = (\hat{\beta}_{11}, \hat{\beta}_{12}, \dots, \hat{\beta}_{1(m+k)}, \dots, \hat{\beta}_{p1}, \hat{\beta}_{p2}, \dots, \hat{\beta}_{p(m+k)})^T$, $\mathbf{B}$ merupakan matriks berorde $n \times (m+k)$, dan $\mathbf{y}$ merupakan vektor dari variabel respon yang berorde $n \times 1$.

Estimasi model fungsi *B-spline* pada regresi nonparametrik diperoleh seperti pada Persamaan (20) berikut:

$$\begin{aligned} \hat{y} &= \mathbf{B}\hat{\boldsymbol{\beta}} \\ &= \mathbf{B}((\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{y}) \\ &= \mathbf{B}(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{y} \\ &= \mathbf{S}\mathbf{y} \end{aligned} \quad (20)$$

dengan $\mathbf{S} = \mathbf{B}(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T$ merupakan matriks *smoother* yang bersifat simetris dan defisit positif.

Selanjutnya, estimasi model regresi nonparametrik *B-spline* dapat dinyatakan seperti Persamaan (21) berikut:

$$\hat{y}_i = \hat{\beta}_{11} B_{1,m}(x_{i1}) + \hat{\beta}_{12} B_{2,m}(x_{i1}) + \dots + \hat{\beta}_{p(m+k)} B_{m+k,m}(x_{ip}) \quad (21)$$

1.5.6 Penentuan Titik Knot Optimal

Pada regresi *spline*, titik knot berfungsi sebagai penanda atau sambungan yang menunjukkan adanya perubahan pola pada interval tertentu. Titik knot juga sering disebut sebagai parameter penghalus karena posisinya menentukan bagaimana kurva *spline* terbentuk dan tersegmentasi. Penentuan titik knot yang optimal tidak hanya bertujuan untuk menghasilkan kurva yang sesuai dengan pola data, tetapi juga untuk menghindari masalah *overfitting* maupun *underfitting*. Oleh karena itu, titik knot biasanya dipilih pada bagian data yang mengalami perubahan kemiringan (*slope*) signifikan agar mampu menggambarkan pola hubungan dengan lebih baik.

Salah satu metode yang umum digunakan untuk menentukan knot optimal adalah *Generalized Cross Validation* (GCV). Model regresi *B-spline* yang baik diperoleh ketika penentuan titik knot didasarkan pada nilai GCV minimum (Handajani ddk., 2023).

$$GCV(K) = \frac{MSE(K)}{\left(\frac{1}{n} \text{trace}[\mathbf{I} - \mathbf{S}(K)] \right)^2}$$

dengan $K = [K_1, K_2, \dots, K_k]$ adalah titik knot, $MSE(K) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$, dan $S(K) = B(B^T B)^{-1} B^T$.

1.5.7 Inflasi

Secara bahasa, inflasi dapat diartikan sebagai kenaikan harga suatu barang yang disebabkan oleh turunnya nilai uang (uang kertas) sebagai akibat dari jumlah uang yang beredar lebih banyak dibandingkan dengan ketersediaan barang. Dari sudut pandang ekonomi, inflasi juga dapat didefinisikan sebagai kecenderungan kenaikan tingkat harga umum secara terus-menerus dalam jangka waktu tertentu yang berakibat pada menurunnya daya beli uang (Pujadi, 2022).

Beberapa karakteristik umum inflasi yang dapat diidentifikasi yaitu:

- d. Kecenderungan meningkat. Meskipun dalam jangka pendek harga beberapa produk tertentu bisa saja turun, namun secara keseluruhan tingkat harga cenderung naik dari waktu ke waktu.
- e. Kontinuitas. Inflasi bukanlah peristiwa sesaat, melainkan proses yang berlangsung terus-menerus dalam suatu periode tertentu.
- f. Bersifat umum. Kenaikan harga tidak hanya terjadi pada satu atau dua jenis barang, melainkan meluas pada sebagian besar barang/jasa yang dikonsumsi masyarakat.

Indeks Harga Konsumen (IHK), yaitu indeks yang menghitung rata-rata perubahan harga dalam suatu periode, merupakan indikator yang umum digunakan untuk mengukur inflasi. Ada dua cara dalam menghitung inflasi, yaitu (Utari dkk., 2016):

1. Inflasi bulanan (*month-to-month*)

Laju inflasi bulanan dihitung berdasarkan selisih indeks harga pada bulan berjalan terhadap indeks pada bulan sebelumnya.

$$\text{inflasi bulan ke } -b = \frac{\text{IHK bulan } b - \text{IHK bulan } (b-1)}{\text{IHK bulan } (b-1)} \times 100\%$$

2. Inflasi tahunan (*year-to-year*)

Laju inflasi tahunan dihitung dengan membandingkan indeks harga pada suatu bulan dengan indeks pada bulan yang sama pada tahun sebelumnya.

$$\text{inflasi bulan ke } -b = \frac{\text{IHK bulan } b \text{ tahun } t - \text{IHK bulan } b \text{ tahun } (t-1)}{\text{IHK bulan } b \text{ tahun } (t-1)} \times 100\%$$

Berdasarkan data dari BPS, tingkat inflasi Indonesia pada Agustus 2025 secara *year-on-year* tercatat sebesar 2,31%, dengan IHK mencapai 108,51. Angka ini sedikit lebih rendah dari inflasi Juli 2025 yang sebesar 2,37%, namun masih berada dalam rentang sasaran inflasi Bank Indonesia, yaitu $2,5\% \pm 1\%$ (BI, 2025). Secara *month-to-month*, Indonesia bahkan mengalami deflasi sebesar 0,08%, sedangkan secara kumulatif sejak awal tahun (*year-to-date*) tercatat inflasi mencapai 1,60%. Kenaikan harga yang paling dominan berasal dari sektor pangan, serta biaya pendidikan pada periode tahun ajaran baru. Secara keseluruhan, pergerakan inflasi Indonesia sepanjang tahun 2025 cenderung stabil dan terkendali, meskipun masih dipengaruhi oleh beberapa faktor eksternal.

BAB II METODE PENELITIAN

2.1 Jenis dan Sumber Data

Penelitian ini menggunakan pendekatan kuantitatif dengan melibatkan pengumpulan dan analisis data berbentuk numerik. Data yang digunakan merupakan data sekunder mengenai inflasi di Indonesia dan variabel-variabel prediktornya selama periode Januari 2015 hingga Desember 2024. Data tersebut diperoleh dari lembaga resmi, yaitu Badan Pusat Statistik (BPS) melalui laman *website* bps.go.id dan Bank Indonesia melalui laman *website* bi.go.id.

2.2 Variabel Penelitian

Pada penelitian ini, variabel respon yang digunakan adalah inflasi, sedangkan variabel prediktor terdiri atas enam variabel. Adapun variabel-variabel yang digunakan dalam penelitian ini disajikan sebagai berikut:

Tabel 1. Variabel penelitian

Variabel	Keterangan	Satuan
Y	Inflasi	%
X_1	Nilai Tukar Rupiah terhadap Dolar Amerika Serikat (Kurs)	Rupiah per 1 USD
X_2	Suku Bunga Acuan	%
X_3	Jumlah Uang Beredar	Miliar Rupiah
X_4	Indeks Harga Pangan	Indeks
X_5	Indeks Harga Perdagangan Besar (IHPB)	Indeks
X_6	Nilai Impor	Juta atau Miliar USD

2.3 Metode Analisis

Metode analisis yang digunakan dalam penelitian ini adalah regresi nonparametrik *B-spline* dengan *Principal Component Analysis*. Adapun langkah-langkah analisis yang diterapkan dalam penelitian ini:

1. Memperoleh dan menyajikan estimasi parameter untuk model regresi nonparametrik *B-spline* dengan PCA.
 - a) Mendefinisikan bentuk awal variabel respon dan variabel prediktor.
 - b) Menerapkan PCA pada variabel prediktor yang mengalami multikolinearitas sehingga diperoleh komponen utama W yang saling bebas.
 - c) Menyusun model regresi nonparametrik *B-spline* terhadap setiap komponen utama yang diperoleh, dengan bentuk:

$$y_i = \sum_{r=1}^q \sum_{l=1}^{m+k} \beta_{rl} B_{l,m}(w_{ir}) + \varepsilon_i$$

di mana $i = 1, 2, \dots, n$ menyatakan observasi, $r = 1, 2, \dots, q$ menyatakan komponen utama, dan $l = 1, 2, \dots, m + k$ menyatakan fungsi basis B-spline, dengan m sebagai orde spline dan k sebagai jumlah titik knot.

- d) Menyajikan model regresi nonparametrik *B-spline* dengan PCA dalam bentuk matriks.

$$\mathbf{y} = \mathbf{B}_w \boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

- e) Mengestimasi parameter $\boldsymbol{\beta}$ menggunakan metode *Ordinary Least Square* (OLS) dengan meminimumkan jumlah kuadrat *error*.

$$S(\boldsymbol{\beta}) = \mathbf{y}^T \mathbf{y} - 2\boldsymbol{\beta}^T \mathbf{B}_w^T \mathbf{y} + \boldsymbol{\beta}^T \mathbf{B}_w^T \mathbf{B}_w \boldsymbol{\beta}$$

- f) Memperoleh estimasi parameter $\boldsymbol{\beta}$ regresi nonparametrik *B-spline* dalam bentuk matriks.

$$\hat{\boldsymbol{\beta}} = (\mathbf{B}_w^T \mathbf{B}_w)^{-1} \mathbf{B}_w^T \mathbf{y}$$

- g) Menyajikan estimasi model regresi nonparametrik *B-spline* dengan PCA.

$$\hat{\mathbf{y}} = \mathbf{B}_w \hat{\boldsymbol{\beta}}$$

2. Membentuk komponen utama sebagai variabel prediktor baru hasil reduksi dari PCA.

- Melakukan eksplorasi data untuk memperoleh gambaran umum karakteristik data melalui nilai statistik dari masing-masing variabel penelitian.
- Melakukan spesifikasi model dengan memanfaatkan *scatter plot* untuk mengidentifikasi pola hubungan antara variabel respon terhadap setiap variabel prediktor.
- Menguji adanya multikolinearitas antarvariabel prediktor menggunakan *Variance Inflation Factor* (VIF).
- Menyajikan matriks korelasi untuk melihat tingkat keeratan hubungan antarvariabel prediktor.
- Menerapkan metode PCA pada variabel prediktor.
- Membentuk komponen utama sebagai variabel prediktor baru yang bebas linear dan merupakan hasil reduksi dari PCA.

3. Membangun model regresi nonparametrik *B-spline* menggunakan komponen utama hasil PCA sebagai variabel prediktor.

- Memodelkan data menggunakan regresi nonparametrik *B-spline* dengan beberapa pilihan orde dan jumlah titik knot.
- Menentukan model regresi nonparametrik *B-spline* terbaik melalui nilai GCV minimum.
- Memperoleh estimasi parameter regresi nonparametrik *B-spline*.
- Menginterpretasikan model regresi nonparametrik *B-spline* dengan PCA.