

Peramalan Kemiskinan di Kabupaten Banyumas Menggunakan Regresi Nonparametrik dengan Pendekatan Kernel Nadaraya Watson *Adjusted*

Novita Eka Chandra¹, Melda Juliza², Muhammad Hafidh Nashrullah³

¹Universitas Jenderal Soedirman, novita.chandra@unsoed.ac.id

²Universitas Jenderal Soedirman, melda.juliza@unsoed.ac.id

³Universitas Islam Darul 'ulum, hafidh@unisda.ac.id

Abstract. Poverty is a serious challenge faced by the Banyumas Regency Government. Although the poverty rate in this region has shown a declining trend for more than a decade, the pattern of decline has not been linear. This study utilizes time series data representing the percentage of the poor population in Banyumas Regency from 2003 to 2024. This research primarily seeks to forecast the poverty rate in 2030 and to differentiate between the performance of two kernel functions, Gaussian and Epanechnikov, which are applied in nonparametric regression using the adjusted Nadaraya-Watson kernel approach. Analysis results suggest that the model performs best when the bandwidth is set at its optimal value of 0,538909 using the Epanechnikov kernel function. Based on the forecast, the poverty rate in 2030 is estimated to be 12,87%. This result indicates the need for well-planned strategies and policies by the Banyumas Regency Government to reduce the poverty rate over the next six years.

Keywords: Forecasting; Poverty; Kernel Nonparametric Regression; Nadaraya-Watson; Adjusted.

Abstrak. Kemiskinan merupakan tantangan serius yang dihadapi oleh Pemerintah Kabupaten Banyumas. Meskipun tingkat kemiskinan di wilayah ini menunjukkan tren penurunan selama lebih dari satu dekade terakhir, pola penurunannya tidak berlangsung secara linier. Penelitian ini menggunakan data deret waktu berupa persentase penduduk miskin di Kabupaten Banyumas dari tahun 2003 hingga 2024. Tujuan utama penelitian adalah meramalkan tingkat kemiskinan pada tahun 2030 serta membandingkan kinerja Gaussian dan Epanechnikov merupakan dua fungsi kernel yang diterapkan dalam regresi nonparametrik menggunakan pendekatan kernel Nadaraya-Watson yang telah disesuaikan (*adjusted*). Hasil analisis menunjukkan bahwa model terbaik diperoleh dengan *bandwidth* optimal sebesar 0,538909 dan menggunakan fungsi kernel Epanechnikov. Berdasarkan hasil peramalan, tingkat kemiskinan pada tahun 2030 diperkirakan sebesar 12,87%. Hasil ini mengindikasikan perlunya strategi dan kebijakan yang terencana oleh Pemerintah Kabupaten Banyumas untuk menurunkan angka kemiskinan dalam enam tahun ke depan.

Kata Kunci: Peramalan; Kemiskinan, Regresi Nonparametrik Kernel, Nadaraya Watson, Adjusted

1 Pendahuluan

Kemiskinan merupakan isu strategis yang masih menjadi perhatian utama di Kabupaten Banyumas. Kemiskinan juga merupakan masalah berkelanjutan di Kabupaten Banyumas. Sebagai salah satu kabupaten dengan jumlah penduduk besar di Jawa Tengah, Banyumas menghadapi tantangan serius dalam upaya peningkatan kesejahteraan masyarakat dan pembangunan. Data yang dihimpun oleh BPS Kabupaten Banyumas mengindikasikan bahwa proporsi penduduk miskin pada tahun 2024 tercatat sebesar 11,95%, yang menempatkan Banyumas melebihi tingkat rata-rata kemiskinan penduduk di wilayah Provinsi Jawa Tengah mencapai 10,47% pada periode yang sama.

Kondisi kemiskinan yang terjadi di Kabupaten Banyumas menunjukkan tren penurunan dalam lebih dari satu dekade terakhir. Namun, tren penurunan tersebut tidak berlangsung secara linier. Misalnya, pada tahun 2020–2021, dampak pandemi COVID-19 memicu peningkatan kembali dalam angka kemiskinan sehingga menyebabkan banyak lapangan pekerjaan informal hilang dan daya beli masyarakat menurun. Fluktuasi semacam ini menunjukkan bahwa faktor waktu (tahun) memengaruhi tingkat kemiskinan dengan cara yang tidak linier.

Selama ini, banyak studi dan perencanaan kebijakan menggunakan pendekatan regresi linier untuk memodelkan hubungan antara waktu dan tingkat kemiskinan [1]. Namun, asumsi linieritas seringkali tidak sesuai dengan kenyataan yang bersifat dinamis dan nonlinier. Oleh karena itu, dibutuhkan pendekatan yang lebih fleksibel untuk menangkap pola perubahan tingkat kemiskinan secara akurat, khususnya dalam konteks data *time series*.

Salah satu pendekatan yang mampu mengakomodasi bentuk hubungan yang tidak liner dan kompleks adalah regresi nonparametrik, khususnya metode kernel Nadaraya-Watson. Metode ini tidak memerlukan asumsi bentuk fungsi tertentu, sehingga dapat menyesuaikan bentuk kurva terhadap data yang tersedia. Dengan kata lain, pendekatan ini memungkinkan pemodelan hubungan yang lebih fleksibel dan realistik berdasarkan pola data yang diamati secara empiris [2].

Regresi nonparametrik kernel Nadaraya-Watson mempertimbangkan bentuk fungsi kernel dan nilai *bandwidth* yang digunakan [3]. Namun, model ini juga memiliki kelemahan, seperti *boundary bias* dan sensitivitas terhadap pemilihan *bandwidth* serta jenis fungsi kernel. Untuk mengatasi hal ini, dapat dilakukan penyesuaian (*adjustment*), baik dalam bentuk *bandwidth* adaptif maupun pemilihan kernel yang lebih sesuai dengan karakteristik data. Regresi nonparametrik kernel *adjusted* dikembangkan oleh [4] dimana fungsi kernel yang digunakan yaitu kernel Gaussian dengan fungsi kernelnya disesuaikan kedalam bentuk skala-lokasi. Dalam papernya dijelaskan regresi nonparametrik kernel Nadaraya Watson *adjusted* lebih baik dibandingkan dengan regresi nonparametrik kernel Nadaraya Watson, yang dilihat dari nilai *error* terkecilnya.

Dengan mempertimbangkan latar belakang tersebut, studi ini memusatkan perhatian pada penerapan regresi nonparametrik dengan pendekatan kernel Nadaraya-Watson *adjusted* untuk meramalkan tingkat kemiskinan di Kabupaten Banyumas. Selain itu, dilakukan perbandingan performa antara dua fungsi kernel berbeda: Gaussian dan Epanechnikov, karena belum banyak studi yang membandingkan efektivitas berbagai jenis kernel dalam konteks data *time series* kemiskinan daerah. Dengan model ini, diharapkan diperoleh gambaran dan

memberikan masukan bagi perumusan kebijakan pengentasan kemiskinan yang lebih tepat sasaran sesuai target.

2 Metode Penelitian

Sumber data dalam penelitian ini berupa data sekunder yang diambil dari publikasi BPS Kabupaten Banyumas. Data ini mencerminkan tingkat kemiskinan penduduk setiap tahunnya (2003-2024) di Kabupaten Banyumas. Data ini berupa data *time series* dimana dengan identifikasi variabel prediktornya (X) adalah tahun dan variabel responnya (Y) adalah persentase penduduk miskin. Pada penelitian ini menggunakan model regresi nonparametrik dengan pendekatan kernel Nadaraya Watson *adjusted*, yang mana fungsi kernel yang digunakan yaitu fungsi kernel Gaussian dan Epanechnikov, dengan bantuan program R. Adapun nilai *bandwidth* yang digunakan yaitu nilai *bandwidth* optimum berdasarkan hasil perhitungan dan menggunakan bantuan *library(KernSmooth)* yaitu fungsi *dpill* pada R [5], [6].

2.1 Regresi Nonparametrik Kernel *Adjusted*

Regresi nonparametrik dengan pendekatan kernel Nadaraya Watson *adjusted* menurut [4] adalah perluasan dari regresi nonparametrik kernel Nadaraya Watson, dengan fungsi kernel yang diambil berdasarkan keluarga skala lokasi dari estimator densitas kernel. Bentuk model regresi nonparametrik dengan pendekatan kernel Nadaraya Watson *adjusted* yaitu

$$\hat{Y}_i = \hat{m}_{hA}(x) = \frac{\frac{\sigma}{n^2 h^2} \sum_{i=1}^n \sum_{j=1}^n K\left(\frac{\sigma x - \sigma X_j + h\theta - hX_j}{h^2}\right) Y_j}{\frac{\sigma}{n^2 h^2} \sum_{j=1}^n \sum_{k=1}^n K\left(\frac{\sigma x - \sigma X_j + h\theta - hX_k}{h^2}\right)}, \quad (1)$$

dengan nilai MSE dari $\hat{m}_{hA}(x)$, yaitu

$$MSE[\hat{m}_{hA}(x)] \approx \frac{\sigma \sigma_1^2(x) \int f^2(u) du}{nhf(x)} + \left(\frac{h^2 \{2f'(x)m'(x) + f(x)m''(x)\}^2 Var(X)}{2\sigma^2 f(x)} \right)^2,$$

dan nilai *bandwidth* optimal dan σ optimalnya adalah

$$h = n^{-1/5} \quad (2)$$

$$\sigma_{opt} = \left(\frac{\{2f'(x)m'(x) + f(x)m''(x)\}^2 Var(X)}{f(x)\sigma_1^2(x) \int f^2(u) du} \right)^{-1/5}. \quad (3)$$

2.2 Fungsi Kernel dan *Bandwidth*

Fungsi kernel menurut [3] adalah sebuah fungsi dengan *bandwidth* dinyatakan dengan

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right),$$

yang mana fungsi kernelnya memiliki sifat berikut:

1. $K(x) \geq 0$, untuk semua nilai $x \in R$
2. $\int K(x) dx = 1$
3. $\int xK(x) dx = 0$
4. $\mu_2(k) = \int x^2 K(x) dx < \infty$
5. $\|K\|_2^2 = \int K^2(x) dx < \infty$.

Berikut ini beberapa macam bentuk fungsi kernel, yaitu:

Tabel 1. Fungsi Kernel

Kernel	$K(x)$
Epanechnikov	$\frac{3}{4}(1-x^2)I(x \leq 1)$
Gaussian	$\frac{\pi}{4} \cos\left(\frac{\pi}{2}x\right)I(x \leq 1)$

Parameter *bandwidth h* merupakan parameter penghalus yang digunakan untuk menentukan tingkat kemulusan dari kurva yang diestimasi. Pemilihan *bandwidth* merupakan hal terpenting dalam metode pendekatan kernel [3] karena besar kecilnya nilai *bandwidth* akan memengaruhi kehalusan estimator. Nilai *bandwidth* yang lebih tinggi cenderung menghasilkan estimasi kurva yang lebih mulus. Sebaliknya, *bandwidth* yang lebih kecil cenderung menghasilkan kurva yang tampak lebih kasar atau tidak rata [7]

2.3 Evaluasi Model

Menurut [8], kriteria untuk menentukan estimator terbaik dalam model regresi nonparametrik menggunakan *Root Mean Square Error* (RMSE), yaitu

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}. \quad (4)$$

Semakin kecil nilai RMSE, maka semakin baik hasil estimasinya.

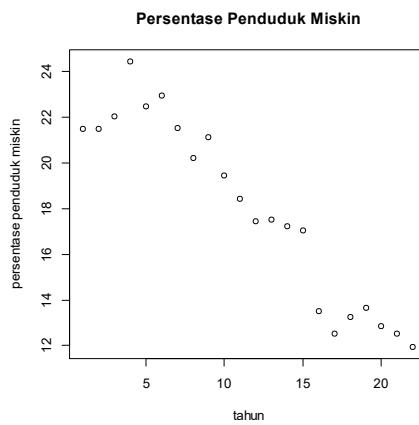
2.4 Rangkaian Proses Penelitian

Adapun rangkaian proses yang dilakukan pada penelitian ini sebagai berikut:

- a. Melakukan plot data dan mendeskripsikan statistika deskriptifnya
- b. Mengestimasi model regresi nonparametrik dengan pendekatan kernel Nadaraya Watson *adjusted* dengan fungsi Gaussian
- c. Mengestimasi model regresi nonparametrik dengan pendekatan kernel Nadaraya Watson *adjusted* dengan fungsi Epanechnikov
- d. Membandingkan hasil estimasi antara fungsi kernel Gaussian, dan Epanechnikov menggunakan *bandwidth* yang optimal pada data kemiskinan, dengan membandingkan plot estimasi kurva regresi bersama-sama dengan plot data serta melihat nilai RMSE
- e. Meramalkan tingkat kemiskinan di Kabupaten Banyumas dengan menggunakan regresi nonparametrik yang terbaik.

3 Hasil dan Pembahasan

Bentuk hubungan antara variabel prediktor (tahun) dengan variabel respon (persentase penduduk miskin) dapat dilihat dari plot antara kedua variabel tersebut (Gambar 1).



Gambar 1. Plot Data Persentase Penduduk Miskin

Gambar 1 menunjukkan bentuk kurva yang menggambarkan hubungan antara tahun dengan persentase penduduk miskin, yang tidak dapat diestimasi menggunakan regresi parametrik. Hal ini dikarenakan bentuk kurva tidak membentuk pola tertentu seperti linear, kuadratik, eksponensial, atau kubik. Pola data yang demikian itu dapat diestimasi menggunakan regresi nonparametrik dengan pendekatan kernel. Penelitian ini memanfaatkan metode kernel Nadaraya Watson yang telah disesuaikan (*adjusted*), supaya hasil estimasi yang diperoleh mengikuti pola data berdasarkan skala lokasinya.

Gambaran umum data yang diolah menggunakan *software* R secara rinci dapat dilihat pada Tabel 2.

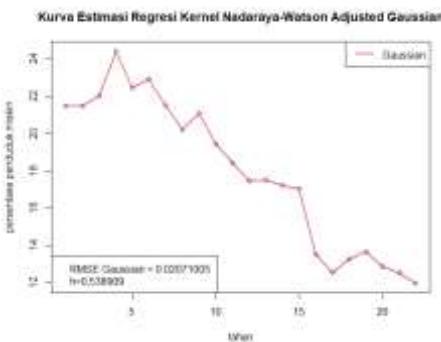
Tabel 2. Statistik Deskriptif Data

Variabel	N	Min	Max	Mean	Standar Deviasi
Tahun	22	1	22	11,5	6,493587
Persentase penduduk miskin	22	11,95	24,44	17,96	4,034975

Jumlah data pengamatan adalah 22, dengan waktu minimum sebesar 1 tahun waktu maksimum sebesar 22 tahun dan persentase penduduk miskin minimum sebesar 11,95%, persentase penduduk miskin maksimum 24,44%. Rata-rata waktu sebesar 11,5 tahun, dan persentase penduduk miskin sebesar 17,69%, serta standar deviasi waktu sebesar 6,493587 tahun dan persentase penduduk miskin sebesar 4,034975%.

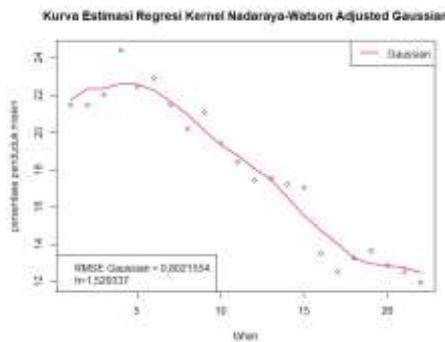
3.1 Model Regresi Nonparametrik dengan Pendekatan Kernel Nadaraya Watson *Adjusted* dengan Fungsi Gaussian

Berikut ini adalah hasil kurva estimasi regresi nonparametrik dengan pendekatan kernel Nadaraya Watson *adjusted* dengan fungsi kernel Gaussian dan *bandwidth* sebesar 0,538909, yang mana nilai *bandwidth* ini diperoleh dari perhitungan pada Persamaan (2), yaitu $h_{opt} = n^{-1/5} = 22^{-1/5} = 0,538909$.



Gambar 2. Kurva Estimasi Regresi Kernel Nadaraya Watson *Adjusted Gaussian* dengan $h=0,538909$

Selanjutnya, jika menggunakan *bandwidth* sebesar 1,529337 (diperoleh dengan *dpill* pada R), maka hasil kurva estimasi regresi nonparametrik dengan pendekatan kernel Nadaraya Watson *adjusted* dengan fungsi kernel Gaussian sebagai berikut.

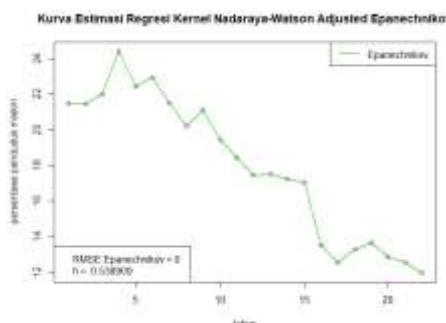


Gambar 3. Kurva Estimasi Regresi Kernel Nadaraya Watson *Adjusted Gaussian* dengan $h=1,529337$

Berdasarkan Gambar 2 dan 3 terlihat bahwa grafik kurva dengan *bandwidth* yang lebih besar memberikan hasil estimasi kurva yang lebih halus, namun tidak tepat pada titik data.

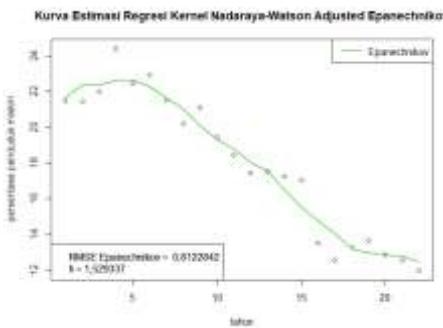
3.2 Model Regresi Nonparametrik dengan Pendekatan Kernel Nadaraya Watson *Adjusted* dengan Fungsi Epanechnikov

Berikut ini adalah hasil kurva estimasi regresi nonparametrik dengan pendekatan kernel Nadaraya Watson *adjusted* dengan fungsi kernel Epanechnikov dan *bandwidth* sebesar 0,538909, yang mana nilai *bandwidth* ini diperoleh dari perhitungan pada Persamaan (2), yaitu $h_{opt} = n^{-1/5} = 22^{-1/5} = 0,538909$.



Gambar 4. Kurva Estimasi Regresi Kernel Nadaraya Watson *Adjusted Epanechnikov* dengan $h=0,538909$

Selanjutnya, jika menggunakan *bandwidth* sebesar 1,529337 (diperoleh dengan *dpill* pada R), maka hasil kurva estimasi regresi nonparametrik dengan pendekatan kernel Nadaraya Watson *adjusted* dengan fungsi kernel Epanechnikov sebagai berikut.



Gambar 5. Kurva Estimasi Regresi Kernel Nadaraya Watson *Adjusted Epanechnikov* dengan $h=1,529337$

Berdasarkan Gambar 4 dan 5 terlihat bahwa grafik kurva dengan *bandwidth* yang lebih besar memberikan hasil estimasi kurva yang lebih halus, namun tidak tepat pada titik data.

3.3 Perbandingan Hasil Estimasi antara Fungsi Kernel Gaussian dan Epanechnikov

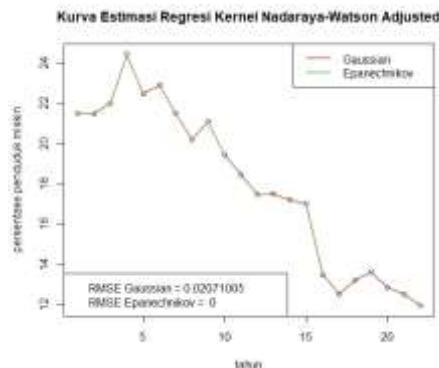
Evaluasi terhadap hasil estimasi dua fungsi kernel, yaitu Gaussian dan Epanechnikov, dalam regresi nonparametrik dilakukan dengan membandingkan nilai RMSE pada Persamaan (4) yang dihasilkan kedua fungsi kernel tersebut yang tercantum pada Tabel 3.

Tabel 3. Perbandingan Hasil Estimasi Kernel Gaussian dan Epanchnikov

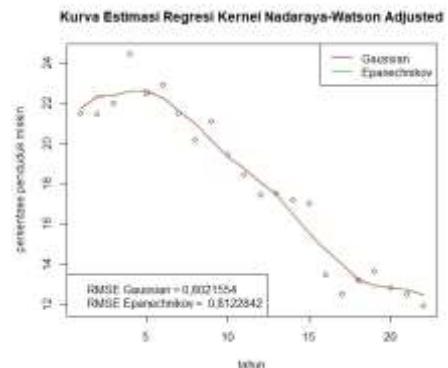
Fungsi Kernel	Bandwidth	RMSE
Gaussian	0,538909	0,02071005
	1,529337	0,8021554
Epanchnikov	0,538909	0
	1,529337	0,8122847

Tabel 3 menunjukkan nilai RMSE yang dihasilkan fungsi kernel Gaussian dan kernel Epanechnikov dengan menggunakan dua nilai *bandwidth* yang berbeda. Secara statistik nilai RMSE yang dihasilkan kernel Epanechnikov lebih kecil dibandingkan kernel Gaussian untuk *bandwidth* terkecil. Namun, jika dilihat dari nilai *bandwidth* yang lebih besar diperoleh bahwa nilai RMSE yang dihasilkan kernel Gaussian lebih kecil dibandingkan kernel Epanechnikov.

Perbandingan selanjutnya, dapat dilihat berdasarkan hasil plot hasil estimasi perbandingan dari kedua fungsi kernel.



Gambar 6. Perbandingan Kurva Estimasi Regresi Kernel Nadaraya Watson *Adjusted* dengan $h=0,538909$



Gambar 7. Perbandingan Kurva Estimasi Regresi Kernel Nadaraya Watson *Adjusted* dengan $h=1,529337$

Pada Gambar 6 dan 7 hasil estimasi kurva regresi antara fungsi kernel Gaussian dan Epanechnikov sangat berimpit dimana menghasilkan bentuk kurva regresi yang sangat mirip. Artinya, penggunaan fungsi kernel yang berbeda dengan *bandwidth* yang optimal yang berbeda untuk masing-masing fungsi kernel tersebut akan menghasilkan estimasi kurva regresi yang sama. Namun, untuk *bandwidth* yang kecil bentuk kurva yang dihasilkan tidak sehalus kurva dengan *bandwidth* besar. Dalam penerapan regresi nonparametrik berbasis kernel Nadaraya Watson *adjusted* tidak hanya berfokus pada besarnya nilai *bandwidth*, tetapi pada fungsi kernel yang sesuai dengan letak skala lokasi titik data [4]. Dengan begitu, estimasi regresi nonparametrik terbaik pada kasus kemiskinan di Kabupaten Banyumas dapat dilihat berdasarkan pada estimasi kurva yang tepat pada titik datanya yaitu dengan nilai *bandwidth* kecil sebesar 0,538909. Untuk fungsi kernel yang sesuai dengan kasus ini yaitu fungsi kernel Epanechnikov dikarenakan menghasilkan nilai RMSE sebesar 0.

3.4 Peramalan

Bentuk model peramalan regresi nonparametrik dengan pendekatan kernel Nadaraya Watson *adjusted* menggunakan fungsi kernel Epanechnikov dan *bandwidth* $h = 0,538909$ pada kasus kemiskinan Kabupaten Banyumas dapat dituliskan sebagai berikut :

$$\hat{Y}_i = \frac{\frac{\sigma}{22^2(0,538909)^2} \sum_{i=1}^{22} \sum_{j=1}^{22} K\left(\frac{\sigma x - \sigma X_i + (0,538909)\theta - (0,538909)X_j}{(0,538909)^2}\right) Y_j}{\frac{\sigma}{22^2(0,538909)^2} \sum_{j=1}^{22} \sum_{k=1}^{22} K\left(\frac{\sigma x - \sigma X_j + (0,538909)\theta - (0,538909)X_k}{(0,538909)^2}\right)}.$$

Apabila nilai parameter skala-lokasi yang telah ditetapkan yaitu nilai θ adalah mean dari tahun dan σ adalah standar deviasi optimum dari tahun yang diperoleh berdasarkan Persamaan (3), maka $\sigma = 0,473158$; $\theta = 11,5$ dan meramalkan persentase penduduk miskin pada tahun 2030 sehingga nilai $X = 28$, diperoleh

$$\hat{Y}_{27} = \frac{\frac{0,473158}{22^2(0,538909)^2} \sum_{i=1}^{22} \sum_{j=1}^{22} K\left(\frac{0,473158(27) - (0,473158)X_i + (0,538909)(11,5) - (0,538909)X_j}{(0,538909)^2}\right) Y_j}{\frac{0,473158}{22^2(0,538909)^2} \sum_{j=1}^{22} \sum_{k=1}^{22} K\left(\frac{(0,473158)(27) - (6,490,473158)X_j + (0,538909)(11,5) - (0,538909)X_k}{(0,538909)^2}\right)}$$

$$\hat{Y}_{27} = 12,87017.$$

Artinya, pada tahun 2030 dengan menggunakan regresi nonparametrik pendekatan kernel Nadaraya Watson *adjusted* maka persentase penduduk miskin di Kabupaten Banyumas sebesar 12,87017% .

4 Kesimpulan

Merujuk pada hasil melalui proses analisis, ditemukan bahwa untuk data persentase kemiskinan di Kabupaten Banyumas diperoleh model regresi nonparametrik dengan pendekatan kernel Nadaraya Watson *adjusted* dengan *bandwidth* optimal sebesar 0,538909 dan fungsi kernel Epanechnikov yang menghasilkan nilai RMSE terkecil. Pada tahun 2030, persentase penduduk miskin di Kabupaten Banyumas menjadi 12,87017%. Berdasarkan hasil peramalan pemerintah daerah perlu melakukan upaya lebih lanjut guna menurunkan tingkat kemiskinan supaya sesuai target dalam 6 tahun ke depan.

5 Daftar Pustaka

- [1] R. Rioke and A. Sugandha, “Analisis Faktor yang Mempengaruhi Kemiskinan di Kabupaten Banyumas,” *Perwira Journal of Science & Engineering*, vol. 2, no. 2, pp. 7–13, 2022.
- [2] M. P. Wand and M. C. Jones, *Kernel Smoothing*. CRC press, 1994.
- [3] W. Hardle, *Applied Nonparametric Regression*. New York: Cambrige, 1994.
- [4] N. E. Chandra, S. Haryatmi, and Z. Zulaela, “Regresi Nonparametrik Kernel Adjusted,” *Jurnal Ilmiah Matematika dan Pendidikan Matematika*, vol. 7, no. 1, pp. 1–10, 2015.
- [5] D. Wang and W. J. Braun, “A Comparison of Bandwidth Selectors for Moderate Degree Local Polynomial Regression,” *Stat Methods Appt*, vol. 34, no. 2, pp. 171–194, 2025.
- [6] F. Fitri and M. Almuhayar, “Comparison of Linear Regression and Polynomial Local Regression in Modeling Prevalence of Stunting,” *Rangkian Mathematics Journal*, vol. 4, no. 1, pp. 16–23, 2025.
- [7] S. A. Putri, A. R. Aristya, N. A. Janad, Y. N. Tadale, and L. Handayani, “Estimator Nadaraya-Watson dengan Fungsi Kernel Normal dan Fungsi Kernel Kuadratik,” *Journal of System and Computer Engineering*, vol. 3, no. 1, pp. 206–219, 2022.
- [8] T. O. Hodson, “Root Mean Square Error (RMSE) or Mean Absolute Error (MAE): When to Use Them or Not,” *Geoscientific Model Development Discussions*, vol. 2022, pp. 1–10, 2022.