

SKRIPSI

**PEMODELAN REGRESI SEMIPARAMETRIK METODE
B-SPLINE**

(Studi Kasus : Indeks Pembangunan Manusia Di Indonesia)



AFA SHOFIATUN NISA'
18106010040
STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA

2025

PEMODELAN REGRESI SEMIPARAMETRIK METODE

B-SPLINE

(Studi Kasus : Indeks Pembangunan Manusia Di Indonesia)

Skripsi

Untuk memenuhi sebagian persyaratan

mencapai derajat Sarjana S-1

Program Studi Matematika



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

diajukan oleh

AFA SHOFIATUN NISA'

18106010040

Kepada

PROGRAM STUDI MATEMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA

YOGYAKARTA

2025



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA



Universitas Islam Negeri Sunan Kalijaga



FM-UINSK-BM-05-03/RO

SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi / Tugas Akhir
Lamp : -

Kepada
Yth. Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga Yogyakarta
di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Afa shofiatun Nisa'
NIM : 18106010040
Judul Skripsi : Pemodelan Regresi Semiparametrik Metode B-spline (Studi kasus:
Indeks Pembangunan Manusia di Indonesia)

sudah dapat diajukan kembali kepada Program Studi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Matematika.

Dengan ini kami mengharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqasyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Yogyakarta, 22 April 2025

Pembimbing

Sri Utami Zuliana, S.Si., M.Sc., Ph.D.

NIP. 19741003 200003 2 002



PENGESAHAN TUGAS AKHIR

Nomor : B-1166/Un.02/DST/PP.00.9/06/2025

Tugas Akhir dengan judul : PEMODELAN REGRESI SEMIPARAMETRIK METODE B-SPLINE (STUDI KASUS : INDEKS PEMBANGUNAN MANUSIA DI INDONESIA)

yang dipersiapkan dan disusun oleh:

Nama : AFA SHOFIATUN NISA`
Nomor Induk Mahasiswa : 18106010040
Telah diujikan pada : Kamis, 12 Juni 2025
Nilai ujian Tugas Akhir : A/B

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang

Sri Utami Zuliana, S.Si., M.Sc., Ph.D.

SIGNED

Valid ID: 6851188ce8a14



Penguji I

Dr. Epha Diana Supandi, S.Si., M.Sc.

SIGNED

Valid ID: 68511da679e40



Penguji II

Arif Munandar, M.Sc.

SIGNED

Valid ID: 685122d5efc8d



Yogyakarta, 12 Juni 2025

UIN Sunan Kalijaga

Dekan Fakultas Sains dan Teknologi

Prof. Dr. Dra. Hj. Khurul Wardati, M.Si.

SIGNED

Valid ID: 68512d6c1138d

SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Afa Shofiatun Nisa'
NIM : 18106010040
Program Studi : Matematika
Fakultas : Sains dan Teknologi

Dengan ini menyatakan bahwa isi skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu Perguruan Tinggi dan sesungguhnya skripsi ini merupakan hasil pekerjaan penulis sendiri sepanjang pengetahuan penulis, bukan duplikasi atau saduran dari karya orang lain kecuali bagian tertentu yang penulis ambil sebagai bahan acuan. Apabila terbukti pernyataan ini tidak benar, sepenuhnya menjadi tanggung jawab penulis.

Yogyakarta, 5 Mei 2025



Afa Shofiatun Nisa'

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSEMBAHAN

Dengan rasa syukur, karya ini dipersembahkan kepada:
Bapak Sholihin dan Ibu Siswati selaku kedua orang tua penulis, dan Adik tercinta Ana Munidlotul Lailiyah, yang selalu memberikan do'a, kasih sayang, serta dukungan. kepada keluarga, teman, dan juga pengasuh Pondok Pesantren Inayatullah yang selalu memberikan dukungan semangat dan dengan segala hormat, kepada Ibu Sri Utami Zuliani selaku dosen pembimbing yang senantiasa sabar membimbing dan mendampingi penulis



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

MOTTO

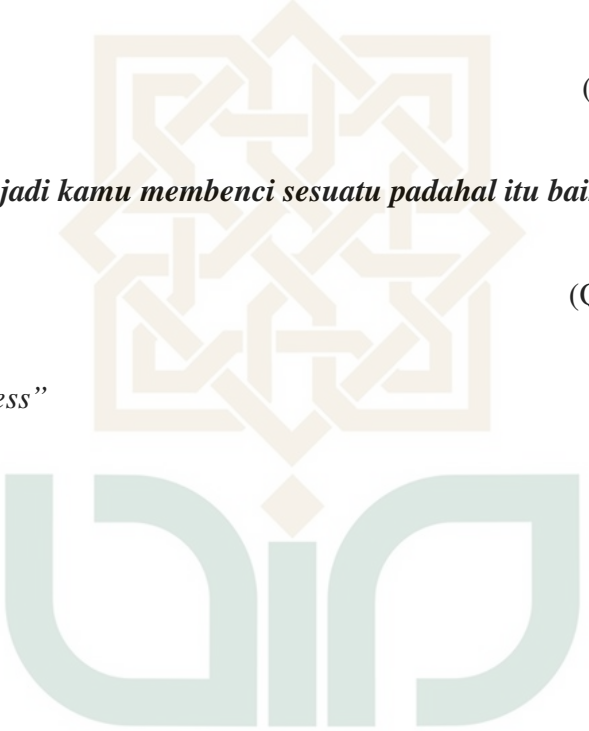
*”Dan katakanlah: Bekerjalah kamu, maka Allah akan melihat pekerja-
anmu.”*

(QS. At-Taubah: 105)

”Boleh jadi kamu membenci sesuatu padahal itu baik bagimu.”

(QS. Al-Baqarah: 216)

”trust the process”



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PRAKATA

Segala puji bagi Allah SWT, Tuhan seluruh alam, atas limpahan rahmat, hidayah, dan karunia-Nya yang telah mengiringi penulis hingga penulis dapat menyelesaikan tugas akhir ini. Shalawat dan salam semoga senantiasa tercurah kepada Nabi Muhammad SAW, beserta keluarga, para sahabat, dan seluruh umat beliau.

Tanpa mengurangi rasa hormat, penulis ingin menyampaikan terimakasih yang sebesar-besarnya kepada:

1. Bapak Prof. Noorhaidi, MA., M.Phil., Ph.D. selaku Rektor UIN Sunan Kalijaga Yogyakarta.
2. Ibu Prof. Dr. Khurul Wardati, M.Si., selaku Dekan Fakultas Sains dan Teknologi, Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
3. Ibu Dr. Epha Diana Supandi, S.Si., M.Sc., selaku Ketua Program Studi Matematika.
4. Bapak Mohammad Farhan Qudratullah, S.Si., M.Si., selaku dosen pembimbing akademik yang telah banyak memberikan pengarahan kepada penulis dari awal perkuliahan hingga akhir.
5. Ibu Sri Utami Zuliana, S.Si., M.Sc., Ph.D. selaku Dosen pembimbing tugas akhir yang selalu sabar membimbing dan memberikan pengarahan kepada penulis dari awal penulisan tugas akhir hingga selesai.
6. Seluruh dosen program studi Matematika dan staf fakultas Sains dan Teknologi yang senantiasa memberikan ilmu dan layanan terbaik kepada penulis

dari awal hingga akhir perkuliahan.

7. Kedua orang tua dan adik penulis yang selalu memberikan kasih sayang, dukungan, serta doa sehingga menjadi alasan penulis menyelesaikan tugas akhir.
8. Bapak Kyai Chamdani Yusuf dan Ibu Nyai Nailil 'Izza selaku pengasuh Pondok Pesantren Inayatullah yang selalu memberi dorongan dan do'a pada penulis.
9. Teman-teman seperjuangan dari Program Studi Matematika angkatan 2018.
10. Teman-teman di Pondok Pesantren Inayatullah yang selalu menemani dan memberikan dukungan serta nasihat-nasihat untuk penulis.
11. serta semua pihak yang tidak dapat penulis sebutkan satu-persatu.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, penulis sangat mengharapkan adanya kritik dan saran yang membangun untuk perbaikan. Penulis berharap karya ini dapat memberikan peran yang positif bagi kemajuan ilmu pengetahuan, khususnya di bidang statistika matematika. Semoga Allah SWT senantiasa memberikan rahmat dan petunjuk-Nya kepada kita semua. *Aamiin.*

Yogyakarta, 22 Maret 2025

Afa Shofiatun Nisa'

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
HALAMAN PERNYATAAN	v
HALAMAN PERSEMBAHAN	vi
HALAMAN MOTTO	vii
PRAKATA	viii
DAFTAR ISI	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMBANG	xv
INTISARI	xvi
ABSTRACT	xvii
I PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian	3
1.4. Manfaat Penelitian	3
1.5. Batasan Masalah	4
1.6. Sistematika Penulisan	4
1.7. Tinjauan Pustaka	5
II DASAR TEORI	8
2.1. Analisis Regresi	8

2.2. Normalisasi Data	9
2.3. Uji Linieritas	10
2.4. <i>Ordinary Least Square</i> (OLS)	10
2.5. Regresi Parametrik	11
2.6. Regresi Nonparametrik	12
2.7. Regresi nonparametrik <i>B-spline</i>	13
2.8. Regresi Semiparametrik	15
2.9. Estimasi Regresi Semiparametrik <i>B-spline</i>	16
2.10. Pemilihan titik knot	18
2.11. Uji Asumsi Residual	19
2.11.1. Uji Asumsi Homogenitas	19
2.11.2. Uji Asumsi Independen	20
2.11.3. Uji Asumsi Normalitas	20
2.12. Uji Multikolinearitas	21
2.13. <i>Cross Validation</i> (CV)	22
2.14. Evaluasi Model	23
2.14.1. <i>Mean Absolute Percentage Error</i> (MAPE)	23
2.14.2. Koefisien Determinasi	24
2.15. Indeks Pembangunan Manusia	24
III METODE PENELITIAN	27
3.1. Jenis Penelitian	27
3.2. Tahapan Penelitian	28
3.3. <i>Flowchart</i> Penelitian	28
IV HASIL DAN PEMBAHASAN	30
4.1. Deskripsi Data	30
4.2. Uji Linieritas	33

4.3. Pemilihan Titik Knot Optimal	36
4.4. Model Regresi Semiparametrik <i>B-spline</i> terbaik	39
4.5. Uji Asumsi Klasik	41
4.5.1. Uji Asumsi Homogenitas	41
4.5.2. Uji Asumsi Independen	42
4.5.3. Uji Asumsi Normalitas	42
4.6. Uji Multikolinearitas	43
4.7. Evaluasi Model	44
4.7.1. MAPE (<i>Mean Absolute Percentage Error</i>)	44
4.7.2. Koefisien Determinasi (R^2)	45
V PENUTUP	47
5.1. Kesimpulan	47
5.2. Saran	49
DAFTAR PUSTAKA	50
LAMPIRAN	57
0.1. Lampiran 1 : Data Penelitian	57
0.2. Lampiran 2: <i>Program Running RStudio</i>	58
0.3. Lampiran 3 : Hasil Uji Asumsi Klasik	72
0.4. Lampiran 4 : Hasil Uji Multikolinearitas	73
0.5. Curriculum Vitae	74

DAFTAR TABEL

2.1	Tabel Kriteria Keputusan Uji Durbin-Watson	20
2.2	Interpretasi Nilai MAPE	23
3.1	Variabel Penelitian	27
4.1	Tabel Statistik Deskriptif	30
4.2	Uji <i>Ramsey</i>	34
4.3	Variabel Data	36
4.4	Nilai GCV berdasarkan Degree dan Knot pada Variabel z_1 dan z_2 . .	37
4.5	Estimasi Parameter Model Regresi Semiparametrik	40
0.1	Hasil Uji Multikolinearitas (GVIF)	73

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR GAMBAR

3.1	flowchart penelitian	29
4.1	Deskriptif variabel x_1	30
4.2	Deskriptif data variabel x_2	31
4.3	Deskriptif data variabel x_3	31
4.4	Deskriptif data variabel x_4	31
4.5	Deskriptif data variabel x_5	31
4.6	<i>scatterplot matrix</i>	33
4.7	<i>plot spline z_1</i>	38
4.8	<i>plot spline z_2</i>	39
4.9	<i>Perbandingan IPM : Aktual vs Prediksi</i>	46

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR LAMBANG

y	= Variabel respon
x_i	= Nilai dari variabel prediktor ke- i
z_i	= Nilai dari variabel prediktor ke- i
$f(z_i)$	= Fungsi nonparametrik ke- i
$B_{j,m}(z)$	= Matrik <i>B-spline</i>
β	= Koefisien regresi
ϵ	= <i>Error</i> acak
Z	= Matriks gabungan parametrik dan B-spline
θ	= Vektor gabungan koefisien yang diestimasi
β_0	= Intersep dari model
α	= Parameter regresi B-spline
m	= orde
k	= titik knot
x_i'	= Data hasil <i>rescalling</i>
$\min(x)$	= Nilai paling rendah dari x
$\max(x)$	= Nilai paling tinggi dari x

INTISARI

PEMODELAN REGRESI SEMIPARAMETRIK METODE B-SPLINE

(Studi Kasus : Indeks Pembangunan Manusia Di Indonesia)

Afa Shofiatun Nisa'

18106010040

Regresi semiparametrik merupakan gabungan dari regresi parametrik dan regresi nonparametrik. Model ini tidak mengharuskan asumsi bentuk fungsi yang spesifik, sehingga memiliki fleksibilitas yang lebih besar dalam menangani pola data yang kompleks dan nonlinier. Fungsi basis *B-spline* digunakan untuk menangani elemen nonparametrik, yang terdiri dari beberapa segmen polinomial dan mampu merepresentasikan bentuk fungsi yang kompleks secara efisien. Penelitian ini bertujuan untuk membangun model regresi semiparametrik berbasis *B-spline* pada data indeks pembangunan manusia. *Mean Absolute Percentage Error* (MAPE) dan Koefisien Determinasi (R^2) digunakan untuk mengetahui keakuratan model. Model regresi semiparametrik *B-spline* terbaik diperoleh pada orde 2, dengan titik knot optimal 0.2517442 untuk variabel (z_1) yang diperoleh dari nilai *Generalized Cross Validation* (GCV) minimum sebesar 0.9055735, dan juga titik knot optimal 0.5913309 untuk variabel (z_2) yang diperoleh dari nilai GCV minimum sebesar 0.4548546. Model tersebut menghasilkan nilai MAPE = 0.3499% yang berarti model yang terbentuk memiliki kemampuan yang sangat akurat untuk menyesuaikan data, dan diperoleh $R^2 = 0.9547$ yang berarti model memiliki kemampuan prediksi yang sangat tinggi terhadap data yang dianalisis.

Kata Kunci: Regresi Semiparametrik, *B-spline*, GCV, Indeks Pembangunan Manusia.

ABSTRACT

SEMIPARAMETRIC REGRESSION MODEL USING B-SPLINE

METHOD

(Case Study: Human Development Index in Indonesia)

Afa Shofiatun Nisa'

18106010040

Semiparametric regression combines parametric regression and nonparametric regression. This model does not require the assumption of a specific function shape, providing greater flexibility in handling complex and nonlinear patterns. B-spline basis functions are used to handle Nonparametric elements consist of several polynomial segments and can efficiently represent complex function shapes. This study aims to develop a semiparametric B-spline regression model using Human Development Index (HDI) data. Mean Absolute Percentage Error (MAPE) and coefficient of determination (R^2) are used to evaluate the model's accuracy. The optimal B-spline semiparametric regression model was obtained at order 2, with an optimal knot point of 0.2450191 for variable z_1 obtained from the minimum Generalized Cross-Validation (GCV) value of 0.9055735, and an optimal knot point of 6,014.621 for variable z_2 obtained from the minimum GCV value of 0.4548546. The model produced an MAPE value of 0.102%, indicating the model's ability to accurately adjust the data. And a R^2 value of 0.9547 indicates its high predictive ability for the analyzed data.

Keywords: Semiparametric Regression, B-spline, GCV, Human Development Index.

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Analisis regresi merupakan salah satu analisis statistika yang digunakan untuk menunjukkan pola hubungan fungsional antara variabel respon dan variabel prediktor (Budiantara, Suryadi, et al., 2006). Variabel prediktor adalah variabel yang digunakan untuk memprediksi atau menjelaskan variabel respon. Variabel respon sendiri adalah variabel yang akan diprediksi atau dijelaskan oleh variabel prediktor (Shofiyah et al., 2018).

Dalam analisis regresi terdapat tiga jenis model pendekatan yang dapat digunakan untuk mempelajari hubungan antar variabel, yaitu Regresi Parametrik, Regresi Non-parametrik, dan Regresi Semiparametrik (Budiantara, 2011). Dalam analisis regresi, apabila pola data sudah diketahui, seperti pola linier, kuadratik, atau kubik, maka metode yang digunakan untuk regresi adalah metode regresi parametrik (Irwan et al., 2024). Namun, dalam praktiknya, tidak semua data mengikuti pola tertentu. Jika pola hubungan antara variabel prediktor dan respon tidak diketahui, maka metode yang tepat digunakan untuk memodelkan hubungan antar variabel tersebut adalah regresi nonparametrik (Budiantara, Lestari, et al., 2010). Selain regresi parametrik dan regresi nonparametrik, ada juga regresi semiparametrik yang merupakan gabungan dari regresi parametrik dan regresi nonparametrik. Dalam regresi semiparametrik, terdapat komponen parametrik dan komponen nonparametrik. Pendekatan ini muncul karena terdapat kasus pemodelan di mana hubungan

antar variabelnya sebagian bersifat linier, sementara sebagian yang lain pola hubungannya tidak diketahui (Budiantara, 2011).

Regresi semiparametrik tidak mengharuskan asumsi bentuk fungsi yang spesifik. Sehingga, regresi semiparametrik memiliki fleksibilitas yang lebih besar dalam menangani pola data yang kompleks dan nonlinier, tanpa terbatas pada regresi parametrik yang cenderung kaku, serta lebih efisien dibandingkan dengan regresi nonparametrik (Khairunnisa et al., 2020). Dalam analisis model regresi semiparametrik, salah satu tantangan yang dihadapi adalah adanya komponen nonparametrik, yang berupa fungsi dengan bentuk yang tidak diketahui. Untuk mengatasi masalah ini, sering digunakan pendekatan nonparametrik yang disebut regresi *spline*. *splines* adalah bentuk polinomial yang tersegmentasi, yang memberikan fleksibilitas lebih dalam menggambarkan fungsi atau karakteristik data. Keunggulan *spline* yaitu pada kemampuannya untuk menggunakan titik knot, yang berfungsi sebagai titik perpaduan bersama dan dapat menunjukkan perubahan pola perilaku data. Dengan menggunakan *spline*, kita dapat menangani heterogenitas dalam data dan memodelkan hubungan kompleks antara variabel tanpa perlu bergantung pada bentuk distribusi yang telah ditentukan sebelumnya. Regresi *spline* menggunakan basis fungsi yang biasanya terdiri dari basis *truncated* dan basis *B-spline* (Eubank, 1999).

Regresi *spline* dengan basis *truncated* memiliki kelemahan yaitu ketika orde *spline* mempunyai derajat yang tinggi, jumlah knot yang banyak, dan jarak antar knot yang terlalu berdekatan. Kondisi ini dapat menyebabkan matriks persamaan normal yang hampir singular, yang mempersulit proses penyelesaiannya. Untuk mengatasi kelemahan ini, digunakan basis *B-spline*, yaitu fungsi basis yang terdiri dari beberapa segmen polinomial. Fungsi basis ini menawarkan fleksibilitas tinggi

dalam menangkap bentuk fungsional yang kompleks dan tidak terbatas (Wulandary et al., 2020).

1.2. Rumusan Masalah

1. Bagaimana model Regresi Semiparametrik yang dihitung menggunakan metode *B-spline*?
2. Bagaimana akurasi model dari Regresi Semiparametrik yang dihitung menggunakan metode *B-spline* pada data Indeks Pembangunan Masyarakat (IPM)?

1.3. Tujuan Penelitian

1. Mengidentifikasi bentuk model Regresi semiparametrik menggunakan metode *B-spline*
2. Mengetahui tingkat keakuratan model Regresi semiparametrik menggunakan metode *B-spline* pada data Indeks Pembangunan Masyarakat (IPM)

1.4. Manfaat Penelitian

1. Memberikan informasi lebih lanjut mengenai penerapan metode regresi semiparametrik, khususnya metode *B-spline*, dalam pemodelan data yang memiliki karakteristik nonlinier pada sebagian variabel prediktor.
2. Menjadi referensi bagi pembaca yang ingin menggunakan metode *B-spline* dalam analisis regresi, khususnya dalam memilih model dengan pendekatan semiparametrik, sebagai alternatif yang fleksibel untuk menangani hubungan nonlinier dalam data.
3. Memberikan informasi terkait tingkat keakuratan model regresi semiparametrik *B-spline* dalam menganalisis dan memodelkan data Indeks Pembangunan

Masyarakat (IPM), yang dapat digunakan sebagai dasar pengambilan keputusan atau kebijakan berbasis data.

1.5. Batasan Masalah

Untuk membatasi masalah agar tidak terjadi perluasan masalah yang baru, akan diaplikasikan sejumlah batasan sebagai berikut:

1. Variabel prediktor dibatasi pada angka harapan hidup, harapan lama sekolah, rata-rata lama sekolah, pengeluaran riil perkapita, dan tingkat pengangguran terbuka Indonesia tahun 2024.
2. Menggunakan metode *Generalized Cross Validation* (GCV) untuk menentukan titik knot yang optimal.
3. Akurasi model dihitung menggunakan *Mean Absolute Percentage Error* (MAPE) dan koefisien determinasi (R^2).
4. Orde yang digunakan adalah orde 2 dan orde 3 dan jumlah titik knot yang digunakan adalah 1 titik knot.

1.6. Sistematika Penulisan

1. BAB I: PENDAHULUAN

Bab ini mencakup latar belakang masalah, rumusan masalah, batasan penelitian, tujuan penelitian, manfaat penelitian, sistematika penulisan, dan tinjauan pustaka.

2. BAB II: DASAR TEORI

Bab ini menjelaskan konsep-konsep teori yang menjadi landasan dalam penelitian, yang meliputi landasan teori, hipotesis, dan rancangan penelitian.

3. BAB III: METODE PENELITIAN

Bab ini menguraikan metode yang digunakan untuk memecahkan masalah penelitian, yang mencakup pendekatan penelitian, sumber data, variabel yang diteliti, tahapan penelitian, serta *flowchart* yang menggambarkan alur penelitian.

4. BAB IV: HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penelitian model regresi semiparametrik dari data index pembangunan manusia menggunakan metode *B-spline*, yang merupakan jawaban dari rumusan masalah dari penelitian ini.

5. BAB V: KESIMPULAN DAN SARAN

Bab ini menyimpulkan hasil pembahasan permasalahan yang telah dibahas, serta memberikan saran terkait penulisan penelitian serupa di masa depan.

1.7. Tinjauan Pustaka

Terdapat beberapa penelitian sebelumnya mengenai metode *B-spline* diantaranya yaitu oleh (Aryansah et al., 2023), penelitian ini menggunakan pendekatan semiparametrik *B-spline* untuk melihat pengaruh harga emas dan minyak mentah dunia terhadap indeks harga saham gabungan. Penelitian ini menggunakan orde 2 dan 4 titik knot untuk mendapatkan model regresi semiparametrik *B-spline* terbaik yang diperoleh dengan mencari nilai GCV terkecil, model *B-spline* terbaik ditemukan dengan nilai koefisien determinasi mencapai 78.8%.

penelitian lain yaitu oleh (Handajani et al., 2023) yang memodelkan produksi padi di Jawa Timur menggunakan regresi nonparametrik *B-spline*. Hasil analisis menunjukkan bahwa suhu, curah hujan, kelembaban, dan luas panen memengaruhi

hasil produksi padi secara signifikan dengan kontribusi R^2 adjusted sebesar 0,987 atau 98.7%.

Penelitian oleh (Sugiantari et al., 2013) menggunakan regresi semiparametrik *spline* untuk menganalisis faktor-faktor yang memengaruhi angka harapan hidup di Jawa Timur. Dalam penelitian ini, pencarian titik optimal dilakukan dengan mencari nilai GCV terendah yang diuji dengan 1 titik knot, 2 titik knot, dan 3 titik knot. Berdasarkan hasil analisis terdapat 3 variabel yang berpengaruh secara signifikan terhadap model, yaitu angka kematian bayi, persentase bayi berusia 0 - 11 bulan yang diberi ASI selama 4-6 bulan, dan persentase balita berusia 1-4 tahun yang diberi imunisasi lengkap, dengan koefisien determinasi sebesar 99.89%.

Penelitian yang dilakukan oleh (Rahmawati et al., 2017) menggunakan regresi nonparametrik *B-spline* untuk memodelkan kasus kemiskinan di Jawa Tengah, dan menganalisis faktor-faktor yang memengaruhinya. Penelitian ini mencoba kombinasi dari orde 2 dan orde 3 dengan banyak titik knot yang akan digunakan untuk setiap orde pada setiap variabel prediktor yaitu satu knot, dua knot dan kombinasi satu dan dua knot untuk mencari model yang terbaik. Berdasarkan hasil analisis, estimasi terbaik diperoleh pada saat x_1 berorde 2, x_2 berorde 2, dan x_3 berorde 2. Dan banyaknya knot x_1 sebanyak 1 knot, x_2 sebanyak 1 knot, dan x_3 sebanyak 2 knot, yang dilihat dari nilai GCV minimum.

Berdasarkan hasil penelitian oleh (Handajani et al., 2023) dan (Rahmawati et al., 2017) yang meneliti tentang regresi nonparametrik *B-spline*, dan penelitian oleh (Sugiantari et al., 2013) dan (Aryansah et al., 2023) yang meneliti tentang regresi semiparametrik *spline* dan *B-spline*, peneliti akan menggunakan regresi semiparametrik *B-spline* untuk memodelkan indeks pembangunan manusia (IPM) di Indonesia dengan faktor angka harapan hidup, harapan lama sekolah, rata-rata la-

ma sekolah, pengeluaran per kapita per tahun, dan tingkat pengangguran terbuka, serta akan mengevaluasi keakuratan model dalam penelitian ini. Berbeda dengan beberapa penelitian sebelumnya yang menggunakan regresi nonparametrik, peneliti akan menggunakan pendekatan semiparametrik dengan metode *B-spline*. Dengan pendekatan semiparametrik, penelitian ini akan memberikan fleksibilitas dalam menangkap pola nonlinear yang kompleks, sekaligus mempertahankan interpretabilitas model. Selain itu, penelitian ini akan mengadopsi metode *Ordinary Least Square* (OLS) sebagai teknik estimasi untuk memperoleh model yang efisien dan mudah diinterpretasikan. Hasil penelitian ini diharapkan dapat memberikan kontribusi signifikan terhadap literatur pembangunan daerah serta memberikan wawasan praktis bagi pengambil kebijakan dan akademisi dalam memahami pola hubungan antara Indeks Pembangunan Manusia (IPM). Selain itu, penelitian ini juga bertujuan untuk menguji validitas eksternal dari metode regresi semiparametrik B-spline pada dataset yang berbeda.

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan hasil penelitian yang dilakukan, beberapa kesimpulan yang dapat diambil adalah sebagai berikut:

1. Model regresi semiparametrik *B-spline* terbaik yang digunakan untuk memodelkan Indeks Pembangunan Manusia (IPM) di Indonesia pada tahun 2024 diperoleh berdasarkan nilai GCV minimum pada orde 2. Titik knot optimal yang diperoleh adalah 0.2517442 untuk variabel angka harapan hidup (z_1) dan 0.5913309 untuk variabel pengeluaran riil per kapita (z_2). Persamaan model regresi semiparametrik terbaik yang terbentuk adalah:

$$\begin{aligned}\hat{y}_i = & 54.081165 + 6.414710 x_{i1} + 8.538050 x_{i2} - 0.026310 x_{i3} \\ & + 2.038034 B_{-1,2}(z_1) + 3.287487 B_{0,2}(z_1) + 5.357792 B_{1,2}(z_1) \\ & + 5.932098 B_{-1,2}(z_2) + 10.641960 B_{0,2}(z_2) + 12.104741 B_{1,2}(z_2)\end{aligned}\quad (5.1)$$

Berdasarkan model yang diperoleh, dapat diketahui bahwa faktor pendidikan seperti variabel harapan lama sekolah (x_1) dan rata-rata lama sekolah (x_2) memiliki pengaruh positif terhadap indeks pembangunan manusia (y) yang masing-masing sebesar 6.414710 dan 8.538050. Setiap tambahan satu tahun pada (x_1) dan (x_2) rata-rata meningkatkan indeks pembangunan manusia secara berarti, menegaskan peran krusial pendidikan dalam mendorong kemajuan pembangunan manusia.

Selain itu, variabel angka harapan hidup (z_1) dan pengeluaran riil per kapita (z_2) yang dimodelkan secara nonlinier melalui basis fungsi *B-spline* juga memberikan

pengaruh positif terhadap indeks pembangunan manusia (y), dengan efek yang semakin kuat pada rentang nilai tertentu. Hal ini memperlihatkan bahwa variabel tersebut tidak bersifat linier, melainkan meningkat secara bertahap sesuai peningkatan variabel.

Sebaliknya, variabel tingkat pengangguran terbuka (x_3) memberikan pengaruh negatif terhadap indeks pembangunan manusia (y) yaitu -0.026310 , hal ini menunjukkan bahwa tingginya tingkat pengangguran terbuka cenderung menurunkan indeks pembangunan manusia.

2. Keakuratan model regresi semiparametrik B-Spline dalam memodelkan indeks pembangunan manusia di Indonesia di evaluasi dengan bantuan MAPE dan koefisien determinasi (R^2).

Dari evaluasi model menggunakan MAPE menghasilkan nilai sebesar :

$$\text{MAPE} = \frac{1}{n} \sum_{j=1}^n \left| \frac{Y_j - \hat{Y}(X_j)}{Y_j} \right| \times 100\% \quad (5.2)$$

$$= 0.3499\% \quad (5.3)$$

Hal ini menunjukkan tingkat kesalahan dari model regresi semiparametrik *B-spline* yang telah dibentuk adalah 0.3499% , Sehingga dapat disimpulkan bahwa model yang terbentuk memiliki kemampuan yang sangat akurat untuk menyesuaikan data.

Evaluasi model berdasarkan koefisien determinasi (R^2) menghasilkan nilai :

$$R^2 = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}$$

$$= 0.9547$$

Nilai R^2 sebesar 0.9547 menunjukkan bahwa sekitar 95.47% dari total variasi dalam indeks pembangunan manusia (y) dapat dijelaskan oleh variabel-variabel prediktor dalam model regresi yang dibangun. Angka R^2 yang sangat tinggi ini menegaskan bahwa model memiliki kemampuan sangat baik dalam memodelkan hubungan antara

variabel-variabel prediktor dan variabel respons, sejalan dengan hasil MAPE yang juga menunjukkan akurasi tinggi.

5.2. Saran

Berdasarkan penelitian yang telah dilakukan, berikut saran yang bisa diberikan untuk penelitian selanjutnya :

1. Disarankan pada penelitian selanjutnya untuk mengeksplorasi variasi jumlah titik knot dan orde yang lebih luas guna memperoleh model yang lebih optimal dan sesuai dengan karakteristik data. Selain itu, penggunaan metode seleksi model yang lain seperti AIC atau BIC dapat dijadikan pembandingan dengan metode GCV untuk mendapatkan hasil yang lebih komprehensif.
2. Sebagai upaya meningkatkan akurasi dan validitas model, disarankan pula untuk melakukan perbandingan antara model B-spline dengan metode spline semiparametrik lain, seperti P-spline (*penalized spline*). Penggunaan P-spline mampu memberikan keleluasaan dalam mengatur tingkat smoothness melalui penalti dan dapat menjadi pembandingan yang efektif untuk menilai keunggulan masing-masing metode dalam konteks data dan masalah yang sama
3. Disarankan agar penelitian selanjutnya mempertimbangkan penambahan variabel independen yang mencerminkan dimensi pembangunan manusia secara lebih luas, seperti derajat urbanisasi, atau indeks ketimpangan. Hal ini bertujuan untuk meningkatkan kemampuan model dalam merepresentasikan dinamika dan keragaman faktor-faktor yang memengaruhi IPM secara menyeluruh.

DAFTAR PUSTAKA

- Adrianingsih, Narita Yuri, Andrea Tri Rian Dani (2021). “Estimasi Model Regresi Semiparametrik Spline Truncated Menggunakan Metode Maximum Likelihood Estimation (MLE)”. In: *Jambura Journal of Probability and Statistics* 2.2, pp. 56–63.
- Ahadi, Giatma Dwijuna, Neni Nur Laili Ersela Zain (2023). “Pemeriksaan Uji Kenormalan dengan Kolmogorov-Smirnov, Anderson-Darling dan Shapiro-Wilk”. In: *Eigen Mathematics Journal*, pp. 11–19.
- Allorerung, Petronilia Palinggik et al. (2024). “Analisis Performa Normalisasi Data untuk Klasifikasi K-Nearest Neighbor pada Dataset Penyakit”. In: *JISKA (Jurnal Informatika Sunan Kalijaga)* 9.3, pp. 178–191.
- Ariesta, Dhea, Nurul Gusriani, Kankan Parmikanti (2021). “Estimasi Parameter Model Regresi Nonparametrik B-Spline E Pada Angka Kematian Maternal”. In: *Jurnal Matematika UNAND* 10.3, pp. 342–354.
- Aryansah, M Pratama, Suparti Suparti (2023). “Pemodelan Regresi Semiparametrik B Spline (Studi Kasus: Pengaruh Harga Emas dan Minyak Mentah Dunia Terhadap Indeks Harga Saham Gabungan)”. In: *Indonesian Journal of Applied Statistics* 6.2, pp. 140–151.
- Azizah, Isna Nur, Prizka Rismawati Arum, Rochdi Wasono (2021). “Model terbaik uji multikolinearitas untuk analisis faktor-faktor yang mempengaruhi produksi padi di Kabupaten Blora tahun 2020”. In: *Prosiding Seminar Nasional UNIMUS*. Vol. 4.

Basuki, Agus Tri (2017). *ANALISIS REGRESI DALAM PENELITIAN EKONOMI DAN BISNIS*.

Budiantara, I Nyoman (2005). “Model Keluarga Spline Polinomial Truncated dalam Regresi Semiparametrik”. In: *BIMIPA* 15.3, pp. 55–61.

— (2011). “Penelitian Bidang Regresi Spline Menuju Terwujudnya Penelitian Statistika yang Mandiri dan Berkarakter”. In: *Prosiding Seminar Nasional MIPA*.

Budiantara, I Nyoman, Fredi Suryadi, et al. (2006). “Pemodelan B-Spline dan MARS Pada Nilai Ujian Masuk terhadap IPK Mahasiswa Jurusan Disain Komunikasi Visual UK. Petra Surabaya”. In: *Jurnal Teknik Industri* 8.1, pp–1.

Budiantara, IN, B Lestari, A Islamiyati (2010). “Estimator Spline Terbobot dalam Regresi Nonparametrik dan Semiparametrik Heteroskedastik untuk Data Longitudinal”. In: *Hibah Penelitian. Jakarta: Kompetensi DP2M-Dikti*.

Budiwanto, Setyo (2017). “Metode statistika untuk mengolah data keolahragaan”. In: *Metode Statistika*, pp. 1–191.

Dani, Andrea Tri Rian, Ludia Ni'matuzzahroh (2022). “Penerapan Keluarga Model Spline Truncated Polinomial pada Regresi Nonparametrik”. In: *Inferensi* 5.1, pp. 37–44.

Dewi, I Gusti Ayu Made Valentina, I Gusti Ayu Made Srinadi, Made Susilawati (2020). “Pemodelan Kasus Pneumonia Pada Balita Di Provinsi Bali Menggunakan Metode Regresi Nonparametrik B-Spline”. In: *E-Jurnal Matematika* 9.3, pp. 197–204.

Dewi, Sarita Permata (2012). “Pengaruh pengendalian internal dan gaya kepemimpinan terhadap kinerja karyawan SPBU Yogyakarta (studi kasus pada spbu anak cabang perusahaan RB. Group)”. PhD thesis. Yogyakarta State University.

- Draper, Norman, Harry Smith (1992). “Analisis Regresi Terapan Edisi II”. In: *Jakarta: PT Gramedia*.
- Erlando, Erlando, Yundari Yundari, Helmi Helmi (2022). “PEMODELAN REGRESI SEMIPARAMETRIK SPLINE TRUNCATED PADA DATA TINGKAT PENGANGGURAN TERBUKA DI INDONESIA”. In: *Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya* 11.3.
- Eubank, Randall L (1999). *Nonparametric regression and spline smoothing*. CRC press.
- Firmansyach, Wildan Attariq, Umi Hayati, Yudhistira Arie Wijaya (2023). “Analisa Terjadinya Overfitting Dan Underfitting Pada Algoritma Naive Bayes Dan Decision Tree Dengan Teknik Cross Validation”. In: *JATI (Jurnal Mahasiswa Teknik Informatika)* 7.1, pp. 262–269.
- Gujarati, Damodar N., Dawn C. Porter, Sukumaran Gunasekar (2012). *Basic Econometrics*. 5th ed. McGraw-Hill.
- Gujarati, DN (2006). “Essential of Econometrics 3rd Edition McGraw–Hill”. In: *New York*.
- Handajani, Sri Sulistijowati et al. (2023). “Pemodelan Produksi Padi di Provinsi Jawa Timur dengan Regresi Non Parametrik B-Spline”. In: *PYTHAGORAS: Jurnal Matematika dan Pendidikan Matematika* 18.2, p. 6.
- Hastie, Trevor et al. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Vol. 2. Springer.
- Hyndman, Rob J., George Athanasopoulos (2018). *Forecasting: Principles and Practice*. 2nd ed. OTexts. URL: <https://otexts.com/fpp2/>.
- Irwan, Muh, Wahyuni Abidin, Wahidah Alwi, et al. (2024). “Model Regresi Nonparametrik dengan Pendekatan Deret Fourier Pada Kasus Tingkat Kemiskinan

- di Sulawesi Selatan”. In: *Jurnal MSA (Matematika dan Statistika serta Aplikasinya)* 12.2, pp. 79–87.
- Khairunnisa, Laili Rahma, Alan Prahutama, Rukun Santoso (2020). “PEMODELAN REGRESI SEMIPARAMETRIK DENGAN PENDEKATAN DERET FOURIER (Studi Kasus: Pengaruh Indeks Dow Jones dan BI Rate Terhadap Indeks Harga Saham Gabungan)”. In: *Jurnal Gaussian* 9.1, pp. 50–63.
- Kuhn, Max, Kjell Johnson (2013). *Applied Predictive Modeling*. Springer. DOI: 10.1007/978-1-4614-6849-3.
- Lase, Novita Ria, Fristi Riandari (2020). “Perancangan Aplikasi Prediksi Jumlah Pendaftar Siswa Baru Dengan Metode Regresi Linier (Studi Kasus: SMA RK Deli Murni Bandar Baru)”. In: *Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)* 3.3.
- Makmur, Zulkifli, Muhammad Arsyam, Andi Muhammad Shaleh Alwi (2020). “Strategi Komunikasi Pembelajaran Di Rumah Dalam Lingkungan Keluarga Masa Pandemi”. In: *KOMUNIDA: Media Komunikasi dan Dakwah* 10.02, pp. 231–241.
- Mardiatmoko, Gun (2020). “Pentingnya uji asumsi klasik pada analisis regresi linier berganda (studi kasus penyusunan persamaan allometrik kenari muda [*canarium indicum* l.]” In: *BAREKENG: Jurnal Ilmu Matematika Dan Terapan* 14.3, pp. 333–342.
- Nabillah, Ida, Indra Ranggadara (2020). “Mean Absolute Percentage Error untuk Evaluasi Hasil Prediksi Komoditas Laut”. In: *Journal of Information System* 5.2, pp. 250–255.
- Naibaho, Mariana Mariana, Ulya Nabila (2021). “Pengaruh produk domestik regional bruto (pdrb) dan tingkat pengangguran terbuka terhadap indeks pema-

- ngunan manusia di kabupaten langkat”. In: *Jurnal Gamma-Pi* 3.2, pp. 21–26.
- Nasution, Darnisa Azzahra, Hidayah Husnul Khotimah, Nurul Chamidah (2019). “Perbandingan normalisasi data untuk klasifikasi wine menggunakan algoritma K-NN”. In: *Comput. Eng. Sci. Syst. J* 4.1, p. 78.
- Pratowo, Nur Isa (2012). “Analisis faktor-faktor yang berpengaruh terhadap Indeks Pembangunan Manusia”. In: *Jurnal Studi Ekonomi Indonesia* 1.1, pp. 15–31.
- Putra, I Made Budiantara, IGAM Srinadi, I Wayan Sumarjaya (2015). “Pemodelan regresi spline”. In: *E-Journal Matematika* 4.3, pp. 110–114.
- Rahmawati, Anisa Septi, Dwi Ispriyanti, Budi Warsito (2017). “Pemodelan Kasus Kemiskinan di Jawa Tengah Menggunakan Regresi Nonparametrik Metode B-Spline”. In: *Jurnal Gaussian* 6.1, pp. 11–20.
- Reza, Widya (2024). *Buku Ajar Analisis Regresi” Pendekatan Praktis dan Sistematis*. Eureka Media Aksara.
- Salam, Nur (2013). “Estimasi Likelihood Maximum Penalized dari Model Regresi Semiparametrik”. In: *Prosiding Seminar Nasional Statistika Universitas Diponegoro 2013*, pp. 571–582.
- Sari, Esa Anindika et al. (2020). “Perbandingan Regresi OLS dan Robust MM-Estimation Dalam Kasus DBD di Indonesia 2018”. In: *Jurnal Education and Development* 8.2, pp. 68–68.
- Setiawan, Mohammad Bhakti, Abdul Hakim (2013). “Indeks pembangunan manusia Indonesia”. In: *Jurnal Economia* 9.1, pp. 18–26.
- Setyaningrum, Kristina Dewanti, Apriani Dorkas Rambu Atahau, Imanuel Madea Sakti (2020). “Analisis z-score dalam mengukur kinerja keuangan untuk

- memprediksi kebangkrutan perusahaan manufaktur pada masa pandemi covid-19". In: *Jurnal Riset Akuntansi Politala* 3.2, pp. 74–87.
- Shofiyah, Fitriyatus, Ayunin Sofro (2018). "Analisis Regresi Linier Multivariat Pada Kandungan Daun Tembakau". In: *MATHunesa: Jurnal Ilmiah Matematika* 6.2.
- Sholikha, Milatus, Made Susilawati, IGAM Srinadi (2019). "Pemodelan Nilai Kurs terhadap Harga Saham pada Data Longitudinal Menggunakan Regresi Non-parametrik Spline Truncated". In: *E-Jurnal Matematika* 8.4, pp. 259–263.
- Sianturi, Rektor (2022). "Uji homogenitas sebagai syarat pengujian analisis". In: *Jurnal Pendidikan, Sains Sosial, Dan Agama* 8.1, pp. 386–397.
- Silalahi, Risda Astridawati et al. (2024). "HASIL PERHITUNGAN ASUMSI KLASIK: TENTANG UJI AUTOKORELASI, NORMALITAS, DAN HETEROKEDATISITAS". In: *Jurnal Ilmiah Multidisipliner* 8.12.
- Siyoto, Sandu, Muhammad Ali Sodik (2015). *Dasar metodologi penelitian*. Literasi Media Publishing.
- Statistik, Badan Pusat (2020). "Indeks pembangunan manusia". In: *Retrieved Februari* 18.
- (2025). *Website Resmi Badan Pusat Statistik*. Diakses pada 12 Maret 2025. URL: <https://www.bps.go.id>.
- Sugiantari, Ayuk Putri, I Nyoman Budiantara (2013). "Analisis Faktor-faktor yang mempengaruhi angka harapan hidup di Jawa Timur menggunakan Regresi Semiparametrik Spline". In: *Jurnal sains dan Seni ITS* 2.1, pp. D37–D41.
- Tediwibawa, Ronald, Desi Yuniarti, Memi Nor Hayati (2019). "Regresi Nonparametrik Spline Birespon Untuk Memodelkan Persentase Penduduk Miskin dan Indeks Kedalaman Kemiskinan di Kalimantan Timur Tahun 2015". In: *EKSPONENSIAL* 10.1, pp. 21–28.

- Wibowo, Wahyu, Sri Haryatmi, I Nyoman Budiantara (2013). “Kajian Metode Estimasi Parameter dalam Regresi Semiparametrik Spline”. In: *BIMIPA* 23.1, pp. 102–110.
- Wijiyanto, Wijiyanto et al. (2024). “Teknik K-Fold Cross Validation untuk Mengevaluasi Kinerja Mahasiswa”. In: *Jurnal Algoritma* 21.1, pp. 239–248.
- Wulandary, Septie, Drajat Indra Purnama (2020). “Perbandingan Regresi Nonparametrik Kernel dan B-Splines pada Pemodelan Rata-Rata Lama Sekolah dan Pengeluaran Perkapita di Indonesia”. In: *Jambura Journal of Probability and Statistics* 1.2, pp. 89–97.
- Wulansari, Andhita Dessy (2016). *Aplikasi Statistika Parametrik dalam Penelitian*.
- Yani, NWMN, I Gusti Ayu Made Srinadi, I Wayan Sumarjaya (2017). “Aplikasi Model Regresi Semiparametrik Spline Truncated”. In: *Jurnal Matematika* 6.1, pp. 65–73.
- Yuniartika, Yuyun, Dadan Kusnandar, M Novitasari Mara (2013). “Penentuan generalized cross validation (gcv) sebagai kriteria dalam pemilihan model regresi b-spline terbaik”. In: *Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya* 2.02.
- Zulfikar, SP, I Nyoman Budiantara (2015). *Manajemen riset dengan pendekatan komputasi statistika*. Deepublish.