

**SKRIPSI**

**OPTIMASI MODEL REGRESI *PENALIZED SPLINE*  
DENGAN METODE *BAYESIAN INFORMATION*  
*CRITERION*(BIC)**

(STUDI KASUS: PRESENTASE INDEKS PEMBANGUNAN MANUSIA (IPM) DI  
KOTA/KABUPATEN PULAU JAWA 2022)



**SHELIEA CANDARA PUSPITA SARI**  
**19106010007**

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

**2023**

**OPTIMASI MODEL REGRESI *PENALIZED SPLINE*  
DENGAN METODE *BAYESIAN INFORMATION*  
*CRITERION*(BIC)**

(STUDI KASUS: PRESENTASE INDEKS PEMBANGUNAN MANUSIA (IPM) DI

KOTA/KABUPATEN PULAU JAWA 2022)

Skripsi

Untuk memenuhi sebagian persyaratan  
mencapai derajat Sarjana S-1  
Program Studi Matematika



diajukan oleh

**SHELIEA CANDARA PUSPITA SARI**

**19106010007**

**Kepada**

**PROGRAM STUDI MATEMATIKA**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA**

**YOGYAKARTA**

**2023**



## **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 : Sheliea Candara Puspita Sari

NIM : 19106010007

Judul Skripsi : Optimasi Model Regresi *Penalized Spline* dengan Metode *Bayesian Information Criterion* (BIC) (Studi Kasus: Indeks Pembangunan Manusia di Kota/Kabupaten Pulau Jawa 2022).

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, 10 Agustus 2023

Pembimbing

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

NIP. 19741003 200003 2 002



KEMENTERIAN AGAMA  
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA  
FAKULTAS SAINS DAN TEKNOLOGI

Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-2293/Un.02/DST/PP.00.9/08/2023

Tugas Akhir dengan judul : OPTIMASI MODEL REGRESI PENALIZED SPLINE DENGAN METODE BAYESIAN INFORMATION CRITERION (BIC) (Studi Kasus: Presentase Pengeluaran per kapita terhadap Indeks Pembangunan Manusia (IPM) di Kota/Kabupaten Pulau Jawa 2022)

yang dipersiapkan dan disusun oleh:

Nama : SHELIEA CANDARA PUSPITA SARI  
Nomor Induk Mahasiswa : 19106010007  
Telah diujikan pada : Selasa, 15 Agustus 2023  
Nilai ujian Tugas Akhir : A-

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: 64e566d29030e



Penguji I

Muhamad Rashif Hilmi, S.Si., M.Sc.  
SIGNED

Valid ID: 64e2d67860418



Penguji II

Aulia Khifah Futhona, M.Sc.  
SIGNED

Valid ID: 64e594798d0a1



Yogyakarta, 15 Agustus 2023  
UIN Sunan Kalijaga  
Dekan Fakultas Sains dan Teknologi

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

Valid ID: 64e6dd3c0f5f7

## SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Sheliea Candara Puspita Sari  
NIM : 19106010007  
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, 11 Agustus 2023



Sheliea Candara Puspita Sari

STATE ISLAMIC UNIVERSITY  
SUNAN KALIJAGA  
YOGYAKARTA

## PERSEMBAHAN

*Karya tulis ini, penulis persembahkan untuk :*

*Ibu tercinta Ibu Listari, kakak dan adik saya Marceliea Fuji Astuti, dan Rosalina Hormat Hayati sebagai rasa hormat, bakti dan ucapan terimakasih yang tidak henti memberikan do'a, motivasi , cinta dan kasih sayang serta dukungan yang tak terhingga.*

*Keluarga, sahabat, teman, dan semua orang yang telah memberikan motivasi secara langsung maupun tidak langsung .*

*Almamater Universitas Islam Negeri Sunan Kalijaga  
Yogyakarta.*

STATE ISLAMIC UNIVERSITY  
SUNAN KALIJAGA  
YOGYAKARTA

## MOTTO

*”Tidak ada yang membatasi mimpi kita kecuali kita sendiri”*

(Nadira Afifah)

*”Sesungguhnya Allah tidak akan mengubah nasib suatu kaum sehingga mereka mengubah keadaan yang ada pada diri mereka sendiri”*

(QS. Ar-Ra’d: 11)



STATE ISLAMIC UNIVERSITY  
SUNAN KALIJAGA  
YOGYAKARTA

## PRAKATA

*Assalamu'alaikum Wr.Wb*

Puji syukur kehadirat Allah SWT yang telah memberikan rahmat, hidayah, dan karunia-Nya berupa keimanan, kesehatan, kekuatan, kesabaran, kelancaran dan kemudahan dalam menyelesaikan skripsi yang berjudul "Optimasi Model Regresi *Penalized Spline* Dengan Metode *Bayesian Information Criterion* (BIC). Sholawat serta salam tetap tercurahkan kepada Nabi Muhammad SAW yang kita nantikan syafaatnya di *yaumul qiyamah*.

Skripsi ini disusun untuk memenuhi persyaratan untuk mendapatkan gelar Sarjana Matematika (S.Mat) pada Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta. Penulis menyadari bahwa dalam penulisan skripsi ini tidak lepas mendapatkan banyak bantuan, motivasi, masukan, dan arahan dari semua pihak. Oleh karena itu penulis ingin mengucapkan rasa hormat dan terima kasih kepada:

1. Ibu Dr. Khurul Wardati, M.Si. selaku Dekan Fakultas Sains dan Teknologi, Universitas Islam Negeri Sunan Kalijaga.
2. Bapak Muchammad Abrori, S.Si., M.Kom. selaku Ketua Program Studi Matematika beserta jajarannya.
3. Bapak Dr. Sugiyanto, S.Si., M.Si., selaku Dosen Pembimbing Akademik yang telah membimbing saya selama perkuliahan.
4. Ibu Sri Utami Zuliana, S.Si., M.Sc., Ph.D. selaku Dosen Pembimbing skripsi yang telah berkenan membimbing, memberikan saran, dan arahan kepada penulis dalam menyelesaikan skripsi ini.
5. Semua tenaga pengajar dan staf di Fakultas Sains dan Teknologi yang senantiasa memberikan pengetahuan dan layanan yang luar biasa sepanjang masa



studi hingga penyelesaian penyusunan skripsi ini.

6. Orang tua tercinta Ibu Listari, Kakak dan adik saya, Kak Marceliea Fuji Astuti dan Dek Rosalina Hormat Hayati yang selalu memberikan dukungan, doa, dan restu terbaik untuk saya.
7. Sahabat seperjuangan saya Viviadita Zahra Putri, Safitri, dan Witarti yang selalu mengingatkan dan memotivasi saya.
8. Sahabat Avenger saya Wida Sivia Fikriana, Rifda Nida'ul Labibah, dan Yuni Pratiwi yang menemani dan mewarnai selama perkuliahan saya.
9. Teman-teman satu bimbingan saya Ara Syaifha Nafia, Devita Riska Nuraini, Gayatri Putri Widowati yang selalu memberikan kritik dan saran untuk mengerjakan skripsi.
10. Teman-teman Matematika angkatan 2019 yang kebersamaian selama 4 Tahun perkuliahan, selalu berbagi ilmu dan pengalaman baik dalam dunia perkuliahan maupun di luar perkuliahan.
11. Terima kasih kepada semua pihak yang tidak mungkin bisa disebutkan satu per satu.

Penulis menyadari akan adanya berbagai kelemahan dan ketidaksempurnaan dalam penyusunan ini. Oleh karena itu, penulis menerima setiap bentuk kritik dan masukan yang diberikan. Penulis berharap agar karya skripsi ini mampu memberikan manfaat, terutama bagi penulis sendiri, serta khalayak umum yang memerlukan informasi ini.

*Wassalamu'alaikum Wr. Wb.*

Yogyakarta, 11 Agustus 2023

Penulis

## DAFTAR ISI

<b>HALAMAN JUDUL</b> . . . . .	i
<b>HALAMAN PERSETUJUAN</b> . . . . .	ii
<b>HALAMAN PENGESAHAN</b> . . . . .	iii
<b>HALAMAN PERNYATAAN</b> . . . . .	iv
<b>HALAMAN PERSEMBAHAN</b> . . . . .	v
<b>HALAMAN MOTTO</b> . . . . .	vi
<b>PRAKATA</b> . . . . .	vii
<b>DAFTAR ISI</b> . . . . .	ix
<b>DAFTAR TABEL</b> . . . . .	xii
<b>DAFTAR GAMBAR</b> . . . . .	xiii
<b>DAFTAR LAMBANG</b> . . . . .	xiv
<b>INTISARI</b> . . . . .	xv
<b>ABSTRACT</b> . . . . .	xvi
<b>I PENDAHULUAN</b> . . . . .	1
1.1. Latar Belakang Masalah . . . . .	1
1.2. Batasan Masalah . . . . .	4
1.3. Rumusan Masalah . . . . .	5
1.4. Tujuan Penelitian . . . . .	5
1.5. Manfaat Penelitian . . . . .	6
1.6. Tinjauan Pustaka . . . . .	6
1.7. Sistematika Penulisan . . . . .	7
<b>II DASAR TEORI</b> . . . . .	9
2.1. Jenis Data Menurut Waktu . . . . .	9

2.2. Analisis Regresi	9
2.3. Analisis Regresi Linier Sederhana	10
2.3.1. Model Regresi Linier Sederhana	11
2.3.2. Estimasi <i>Maximum Likelihood</i>	11
2.3.3. Uji Asumsi	13
2.4. Analisis Regresi Parametrik	14
2.5. Analisis Regresi Nonparametrik	14
2.6. Regresi <i>Spline</i>	15
2.7. Regresi <i>B-Spline</i>	15
2.8. Regresi <i>Penalized Spline</i>	17
2.9. Orde, Knot, dan Parameter <i>Smoothing Optimal</i>	19
2.10. Pengujian Signifikansi Parameter	20
2.11. Pengujian Estimator yang Baik	21
2.12. Indeks Pembangunan Manusia	22
<b>III METODOLOGI PENELITIAN</b>	<b>24</b>
3.1. Jenis dan Sumber Data	24
3.2. Metode Pengumpulan Data	24
3.3. Populasi dan Sampel	24
3.4. Variabel Penelitian	25
3.5. Pengolahan Data	25
3.6. <i>Source Code</i>	25
<b>IV HASIL DAN PEMBAHASAN</b>	<b>26</b>
4.1. Deskripsi Data	26
4.2. Menentukan Orde, Jumlah Knot, Parameter <i>Smoothing Optimal</i>	29
4.3. Interpretasi Model	31
4.4. Pengujian Signifikansi Parameter	33

4.5. Pengujian Kelayakan Model	34
<b>V PENUTUP</b>	<b>36</b>
5.1. Kesimpulan	36
5.2. Saran	37
<b>DAFTAR PUSTAKA</b>	<b>38</b>
<b>LAMPIRAN 1</b>	<b>40</b>
<b>LAMPIRAN 2</b>	<b>43</b>
<b>LAMPIRAN 3</b>	<b>57</b>
<b>Curriculum Vitae</b>	<b>59</b>

## DAFTAR TABEL

2.1 Analisis Variansi Signifikansi Parameter . . . . .	21
4.1 Karakteristik Data IPM di Kota/Kabupaten Pulau Jawa 2022 . . . . .	26
4.2 Orde, Jumlah Knot, dan $\lambda$ Optimal . . . . .	30
4.3 Estimasi Parameter Menggunakan Metode BIC . . . . .	31
4.4 Analisis Uji Serentak Parameter Model BIC . . . . .	33
4.5 Uji Kinerja Model Menggunakan Nilai <i>Mean Percentage Error</i> (MA- PE) . . . . .	34
4.6 Prediksi Pengeluaran per Kapita terhadap IPM di Provinsi Kali- mantan Timur 2022 . . . . .	35

## DAFTAR GAMBAR

4.1 Matriks Korelasi IPM dan Pengeluaran per Kapita . . . . .	28
4.2 Hubungan $\lambda$ dengan nilai BIC . . . . .	30



## DAFTAR LAMBANG

$y$	: Variabel Dependen
$x$	: Variabel Independen
$h(x)$	: Fungsi Kurva Regresi
$\varepsilon$	: Error
$\beta$	: Parameter
$r$	: Titik Knot
$t$	: Jumlah Knot
$s$	: Letak Titik Knot
$v$	: Orde
$u$	: Jumlah Orde
$p$	: Banyak Data
$a$	: Nilai Minimum Variabel Independen
$b$	: Nilai Maksimum Variabel Independen
$Z$	: Basis <i>B-Spline</i>
$\beta$	: Koefisien <i>B-Spline</i>
$\Delta_v$	: <i>Different Operator</i>
$\lambda$	: Parameter Penghalus
$D$	: Matriks Diagonal
$L$	: Fungsi Likelihood
$I$	: Matriks Identitas

## INTISARI

### Optimasi Model Regresi *Penalized Spline* Dengan Metode *Bayesian Information Criterion*(BIC)

(Studi Kasus: Presentase Indeks Pembangunan Manusia (IPM) di Kota/Kabupaten Pulau Jawa  
2022)

Oleh

SHELIEA CANDARA PUSPITA SARI

19106010007

Regresi *penalized spline* merupakan gabungan regresi dengan basis *B-Spline* pada penalti kekasaran diskrit untuk memperhalus kurva. *Penalized spline* mencoba untuk mencapai keseimbangan antara kesesuaian model dengan data dan kehalusan kurva. Dengan memodifikasi koefisien basis B-spline melalui penalti kekasaran diskrit, metode ini dapat menghasilkan kurva yang lebih halus, mencegah overfitting, dan memberikan fleksibilitas dalam menangkap pola dalam data. Pada pemodelan *penalized spline* perlu diperhatikan tingkat kompleksitas model (orde), titik knot, serta parameter penghalus ( $\lambda$ ) yang optimal. Tujuan penelitian ini adalah untuk mengestimasi regresi *penalized spline* pada data IPM. Model regresi yang optimal diperoleh berdasarkan kombinasi pemilihan orde, knot, dan parameter penghalus yang optimal menggunakan nilai minimum *Bayesian Information Criterion* (BIC). Hasil penelitian ini menunjukan model optimal pada orde 2, knot 1, dan lambda ( $\lambda$ ) 0.001 dengan nilai BIC sebesar 543.3806. Estimasi parameter tersebut signifikan dengan nilai MAPE sebesar 0.004059336. Model regresi *penalized spline* optimal yang diperoleh yaitu  $\hat{y} = 8173.879Z_{-1,2}(x) + 8823.742Z_{0,2}(x) + 24217.31Z_{-1,2}(x)$ .

**Kata Kunci:**BIC, Indeks Pembangunan Manusia, MAPE, *Penalized Spline*



## ABSTRACT

### **Optimasi Model Regresi *Penalized Spline* Dengan Metode *Bayesian Information Criterion*(BIC)**

By

SHELIEA CANDARA PUSPITA SARI

19106010007

*Penalized spline regression is a combined regression with a base of B-Spline on a discrete roughness penalty to smooth the curve. It tries to achieve a balance between the fit of the model to the data and the smoothness of the curve. By modifying the B-spline basis coefficients through discrete roughness penalty, this method can produce smoother curves, prevent overfitting, and provide flexibility in capturing patterns in the data. In the modeling of penalized spline, it is necessary to consider the optimal model complexity level (order), knot points, and smoothing parameters ( $\lambda$ ). The purpose of this study is to estimate penalized spline regression on HDI data. The optimal regression model is obtained based on a combination of optimal selection of order, knots, and smoothing parameters using the minimum Bayesian Information Criterion (BIC) value. The results of this study indicate the optimal model at order 2, knot 1, and lambda 0.001 with a BIC value of 543.3806. The parameter estimate is significant with a MAPE value of 0.004059336. The optimal spline regression model obtained is  $\hat{y} = 8173.879Z_{-1,2}(x) + 8823.742Z_{0,2}(x) + 24217.31Z_{-1,2}(x)$ .*

**Keyword:** *BIC, Human Development Index, MAPE, Penalized Spline*

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang Masalah

Statistika merupakan disiplin ilmu yang mempelajari metode pengumpulan, manajemen, penyajian, analisis data, dan pengambilan keputusan pada kondisi tidak ada unsur ketidakpastian. Berdasarkan penggunaannya, statistika dibagi menjadi dua yaitu statistika deskriptif dan statistika inferensia. Statistika yang mempelajari bagaimana cara mengumpulkan, mengolah, menyajikan, dan menganalisis data disebut statistika deskriptif (Supandi, 2020). Analisis data yang termasuk statistika deskriptif antara lain mean, median, modus, pengukuran variansi data, grafik, tabel, dan lainnya yang bertujuan untuk mengetahui kecenderungan data tanpa adanya generalisasi. Sedangkan Statistika inferensial merupakan cabang statistika yang fokus pada metode-metode pengambilan kesimpulan mengenai sifat-sifat keseluruhan populasi berdasarkan data berbentuk angka dari sampel, dengan tujuan melakukan generalisasi atau inferensi.

Statistika inferensia digunakan sebagai penguji perbandingan, pembentukan model, dan melihat korelasi. Dalam pengujian perbandingan, hal yang dapat dibandingkan ukuran data seperti proporsi, varian, dan nilai rata-rata data. Untuk melihat hubungan data dapat dilakukan dengan melihat hubungan sebab akibat. Salah satu metode yang paling sering digunakan adalah analisis regresi. Analisis regresi menjelaskan pengaruh variabel independen terhadap variabel dependen.

Analisis regresi dibagi menjadi tiga yaitu analisis regresi parametrik, analisis regresi semiparametrik, dan analisis regresi nonparametrik. Analisis regresi semiparametrik merupakan gabungan dari analisis regresi parametrik dan nonparametrik. Regresi parametrik digunakan apabila bentuk kurva fungsi regresi diketahui. Hal ini untuk menjelaskan hubungan kedua variabel diperlukan sebuah kurva regresi

yang membentuk pola hubungan. Regresi nonparametrik digunakan untuk memahami hubungan antara variabel independen dan variabel dependen, ketika bentuk kurva fungsi regresi tidak diketahui sebelumnya dari data yang sedang dianalisis. Dalam regresi non-parametrik, bentuk kurva tersebut tidak mengikuti pola tertentu, sehingga metode ini menonjolkan fleksibilitas yang lebih besar.

Model regresi dihasilkan melalui proses estimasi parameter dengan menggunakan metode khusus, seperti metode *maximum likelihood*. Dalam penggunaan *maximum likelihood*, penting untuk memastikan bahwa asumsi-asumsi terpenuhi agar hasil estimasi memenuhi sifat *Best Linear Unbiased Estimator* (BLUE). Namun, metode *maximum likelihood* sangat sensitif terhadap ketidaksesuaian asumsi dalam data. Dalam konteks penerapan regresi parametrik, sering kali terjadi ketidaksesuaian dengan asumsi-asumsi tersebut. Jika salah satu asumsi tidak terpenuhi, estimasi *maximum likelihood* yang dihasilkan dapat kehilangan efisiensi. Sebagai alternatif, regresi nonparametrik digunakan. Dalam regresi nonparametrik, estimasi dilakukan berdasarkan data pengamatan dengan berbagai metode penyalarsan (*smooth*), termasuk penghalus estimator *spline*, estimasi deret ortogonal, estimasi kernel, dan sejenisnya (Eubank, 1999).

Regresi *spline* merupakan salah satu metode yang sering dipakai pada regresi nonparametrik. Regresi *spline* digunakan untuk suatu pendekatan pada plot data dengan memperhitungkan kemulusan kurva. *Spline* kurva dibentuk berdasarkan orde dan titik-titik *knot* yang kemudian saling dihubungkan. Dimana orde dari *spline* merujuk pada derajat polinomial yang digunakan dalam setiap subinterval kurva *spline* dan titik *knot* berarti titik-titik tempat dimana subinterval-subinterval bertemu. Derajat polinomial mengacu pada tingkat tertinggi dari suatu persamaan polinomial.

Regresi *spline* memiliki basis umum yaitu basis fungsi *truncated* dan basis *B-spline*. Regresi *spline* dengan basis fungsi *truncated* memiliki kekurangan ketika orde tinggi maka jumlah *knot* akan bertambah, sehingga letak titik *knot* akan berdekatan dan menciptakan matrik yang hampir singular. Untuk mengatasi masalah ter-

sebut D.Boor menyarankan menggunakan *B-spline*. Namun *B-spline* juga dianggap sensitif terhadap jumlah knot. Jika jumlah knot terlalu banyak akan menghasilkan model *overfit* dan jumlah knot terlalu sedikit mengakibatkan *underfit*. Oleh karena itu Eilers & Marx (1996) menyarankan untuk penambahan penalti atau parameter penghalus dengan menggunakan regresi *penalized spline*. Pada regresi *penalized spline* orde, knot, dan parameter penghalus yang optimal dapat diperoleh menggunakan beberapa metode salah satunya *Bayesian Information Criterion* (BIC). BIC adalah ukuran kebaikan model statistik, dengan mempertimbangkan jumlah parameter dalam model. Parameter smoothing yang optimal adalah yang memberikan nilai BIC terkecil (Durbán et al., 2002).

Penerapan analisis regresi nonparametrik *Penalized spline* pada penelitian ini, menggunakan studi kasus data Indeks Pembangunan Manusia (IPM) di Kota/Kabupaten Pulau Jawa tahun 2022 yang berupa data *cross section*. Data *cross section* merupakan data yang dikumpulkan pada periode waktu tertentu untuk memberikan gambaran keadaan suatu subjek. Data *cross section* ini berbeda dengan data longitudinal yang pengumpulan datanya dalam beberapa titik waktu.

Pada tahun 2020, *World Health Organization* (WHO) menyatakan bahwa Covid-19 mulai menyebar di wilayah Indonesia. Pandemi Covid-19 ini memiliki dampak yang signifikan terhadap Indeks Pembangunan Manusia (IPM) khususnya di Indonesia. Sebelum pandemi Covid-19, capaian IPM mengalami peningkatan sebesar 0,74% dibandingkan tahun sebelumnya. Namun, pada tahun 2020 IPM hanya mengalami sedikit kenaikan sebesar 0,03%. IPM dipengaruhi oleh beberapa faktor salah satunya standar hidup layak diwakili oleh Pengeluaran per Kapita yang disesuaikan. Capaian IPM tahun 2022 merupakan capaian tertinggi bahkan jauh melebihi capaian sebelum pandemi dengan pertumbuhan sebesar 2,90 persen.

Tahun 2022 situasi pandemi Covid-19 semakin terkendali. Sejalan dengan kondisi yang membaik, pada tahun 2022 IPM Indonesia mencapai 72,91, tumbuh 0,86%, lebih cepat dibanding tahun sebelumnya yang hanya 0,49% pada tahun 2021. IPM tertinggi tahun 2022 berada pada provinsi DKI Jakarta dan diikuti be-

berapa kota/kabupaten yang berada di Pulau Jawa seperti D.I.Yogyakarta, Banten, dan Jawa Barat. IPM atau Indeks Pembangunan Manusia, berfungsi sebagai alat untuk menilai pencapaian suatu daerah dalam hal pembangunan manusia. Konsep pembangunan manusia mengacu pada peningkatan kemampuan melalui peningkatan kesehatan, pengetahuan, dan keterampilan individu, serta efektifitas penggunaan keterampilan tersebut. Upaya meningkatkan pembangunan manusia memiliki dampak penting dalam mengurangi tingkat kemiskinan dan kesenjangan, karena akses terhadap pendidikan dan kesehatan yang kuat memberikan kesempatan bagi penduduk yang kurang mampu untuk meningkatkan aset mereka.

Indeks Pembangunan Manusia (IPM) mencerminkan tingkat kemajuan sumber daya manusia di suatu wilayah. Oleh karena itu, berfungsi sebagai patokan untuk mengukur tingkat kesejahteraan di daerah atau negara tertentu. IPM menggabungkan sejumlah indikator yang mencakup aspek-aspek tertentu, diantaranya umur harapan hidup saat lahir yang mencerminkan kesehatan, durasi rata-rata pendidikan dan perkiraan lama sekolah yang mewakili bidang pendidikan, serta indikator ekonomi yang ditunjukkan oleh pengeluaran per kapita yang diatur untuk memperhitungkan indeks daya beli penduduk (BPS, 2022).

Berdasarkan uraian di atas, peneliti tertarik melakukan penelitian mengenai Indeks Pembangunan Manusia (IPM) di Kota/Kabupaten Pulau Jawa pada tahun 2022. Data tersebut dapat diakses pada laman resmi Badan Pusat Statistik melalui [www.bps.go.id](http://www.bps.go.id). Analisis yang digunakan yaitu analisis regresi nonparametrik *penalized spline*. Selanjutnya akan dicari dengan optimasi *Bayesian Information Criterion* (BIC) menggunakan *software R*.

## 1.2. Batasan Masalah

Dalam penyusunan tugas akhir ini, penerapan batasan masalah menjadi sangat penting guna memastikan pembahasan tetap terfokus pada permasalahan yang tengah dibahas, dan menghindari penyimpangan dari tujuan awal. Dengan merujuk pada konteks yang telah dijelaskan sebelumnya, maka batasan masalah dalam tugas

akhir ini diuraikan sebagai berikut:

1. Pemilihan optimasi model regresi nonparametrik *penalized spline* menggunakan *Bayesian Information Criterion* (BIC)
2. Orde yang digunakan orde 2 sampai dengan 4 dan jumlah knot 1 sampai 3.
3. Parameter penghalus ( $\lambda$ ) yang digunakan mulai 0.001 sampai dengan 0.99.
4. Pengolahan data menggunakan bahasa pemrograman R versi 4.2.2.

### 1.3. Rumusan Masalah

Dari dasar yang telah diuraikan, penulis mengidentifikasi beberapa permasalahan dalam penelitian ini, yang dapat dirumuskan sebagai berikut:

1. Bagaimana proses analisis regresi nonparametrik *penalized spline*?
2. Bagaimana penerapan regresi nonparametrik *penalized spline* dalam memodelkan Indeks Pembangunan Manusia (IPM) di Kota/Kabupaten Pulau Jawa pada tahun 2022 ?
3. Bagaimana optimasi model *penalized spline* dengan *Bayesian Information Criterion* (BIC)?

### 1.4. Tujuan Penelitian

Tujuan penulis melakukan penelitian ini, adalah sebagai berikut:

1. Mengkaji tentang analisis regresi nonparametrik *penalized spline*.
2. Mengkaji optimasi *Bayesian Information Criterion* pada regresi *penalized spline*.
3. Pengaplikasian model regresi *penalized spline* pada data Indeks Pembangunan Manusia di Kota/Kabupaten Pulau Jawa pada tahun 2022.

### 1.5. Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini, adalah sebagai berikut:

1. Bagi penulis sebagai tambahan informasi dan wawasan mengenai *penalized spline* dan estimator regresi nonparametrik *Bayesian Information Criterion* (BIC).
2. Bagi para pembaca, dapat menambah pengetahuan mengenai pengaplikasian ilmu statistika khususnya analisis regresi nonparametrik yang dapat dikembangkan lebih lanjut.

### 1.6. Tinjauan Pustaka

Analisis literatur pada penelitian ini mencakup telaah dari berbagai penelitian yang memiliki relevansi dengan topik yang dibahas. Beberapa penelitian yang memiliki kaitan dengan tema yang diangkat oleh penulis adalah sebagai berikut:

Penelitian [Agustina et al. \(2015\)](#) membahas mengenai *Pemodelan Data Indeks Harga Saham Gabungan Menggunakan Regresi Penalized Spline*. Pada penelitian ini menggunakan nilai *Generalized Cross Validation* (GCV) minimum untuk menentukan parameter penghalus  $\lambda$  optimal, sedangkan untuk ketepatan atau keakuratan model menggunakan *Mean Absolute Percentage Error* (MAPE) dan nilai koefisien determinasi ( $R^2$ ).

Penelitian [Zia et al. \(2017\)](#) membahas mengenai *Pemodelan Regresi Spline Menggunakan Metode Penalized Spline Pada Data Longitudinal (Studi Kasus : Harga Penutupan Saham LQ45 Sektor Keuangan dengan Kurs USD terhadap Rupiah Periode Januari 2011-Januari 2016)*. Data yang digunakan adalah data sekunder yang diambil dari saham BRI, BCA, Mandiri dan sebanyak 61 data pengamatan in sample. Pada penelitian ini untuk menentukan parameter smoothing ( $\lambda$ ) dan titik knot menggunakan metode *Generalized Cross Validation* (GCV), sedangkan untuk memilih model terbaik menggunakan koefisien determinasi ( $R^2$ ) dan *Mean Absolute Percentage Error* (MAPE). Pengolahan dan analisis data menggunakan program

R.

Dalam penelitian [Kurniasari et al. \(2019\)](#) menulis penelitian dengan judul "Estimasi Parameter Regresi Spline Dengan Metode *Penalized Spline*" pada data Indeks Pembangunan Manusia menurut Provinsi di Indonesia Tahun 2015. Berdasarkan penelitian tersebut diketahui bahwa model terbaik pada setiap variabel yaitu model *spline* linier dengan nilai GCV minimum. Dihasilkan estimasi parameter yang signifikan dari metode *penalized spline* dan nilai koefisien determinan terkorrelasi sebesar 76,66% serta nilai MAPE sebesar 1,145%.

Pada penelitian [Asikin et al. \(2020\)](#) yang memodelkan Balita *stunting* di Jawa Barat menggunakan *Generalized Additive Models Penalized Spline* diperoleh hasil bahwa pola hubungan variabel respon dan variabel prediktor tidak mudah dispesifikasikan sehingga menggunakan *Generalized Additive Models Penalized Spline* untuk memperoleh distribusi optimal pada variabel respon.

Kesamaan antara penelitian ini dengan studi sebelumnya adalah pendekatan dengan model *Penalized Spline*. Namun, perbedaan yang menonjol adalah pendekatan optimasi yang dipilih, dimana penelitian ini menerapkan *Bayesian Information Criterion (BIC)*, dan berfokus pada kasus Indeks Pembangunan Manusia di Kota/Kabupaten di Pulau Jawa tahun 2022.

### 1.7. Sistematika Penulisan

Dalam rangka memudahkan pemahaman menyeluruh terhadap penelitian ini, penulis akan merinci susunan penulisan sebagai berikut:

#### BAB I PENDAHULUAN:

Bab ini akan mengulas mengenai konteks latar belakang masalah, batasan masalah rumusan masalah, tujuan penelitian, manfaat penelitian, tinjauan pustaka, dan sistematika penulisan.

#### BAB II : DASAR TEORI

Bab ini menjelaskan tentang konsep-konsep teori yang menunjang penelitian sebagai pendukung pembahasan analisis penelitian mengenai Indeks Pembangunan



Manusia di Kota/Kabupaten Pulau Jawa tahun 2022.

### BAB III : METODE PENELITIAN

Pada bab ini membahas mengenai metode penelitian yang digunakan dalam memecahkan masalah, di mana berisi tentang jenis data, metode pengumpulan data, variabel penelitian, dan metode analisis data.

### BAB IV : PEMBAHASAN

Pada bab ini membahas hasil penelitian dari pembahasan permasalahan.

### BAB V : KESIMPULAN DAN SARAN

Bab ini memuat ringkasan dari analisis terhadap isu-isu yang telah dibahas serta rekomendasi yang berkaitan dengan penulisan serupa di masa depan.

## BAB V

### PENUTUP

Pada bab ini akan diberikan kesimpulan dan saran berdasarkan materi yang telah dibahas pada bab sebelumnya.

#### 5.1. Kesimpulan

1. Dalam melakukan analisis untuk memodelkan data menggunakan regresi *penalized spline*, sejumlah langkah dilakukan dalam penelitian ini:
  - a. Mengestimasi model regresi *penalized spline* berdasarkan pilihan orde dan jumlah titik knot.
  - b. Menentukan kombinasi orde dan jumlah titik knot serta parameter penghalus ( $\lambda$ ) yang optimal melalui kriteria BIC minimum.
  - c. Melakukan estimasi parameter terbaik pada model regresi *P-Spline* dengan mempertimbangkan nilai BIC.
  - d. Melakukan pengujian kesignifikanan parameter terhadap model dengan menggunakan uji serentak, dan mengevaluasi kesesuaian model berdasarkan perhitungan nilai MAPE.
2. Analisis regresi *penalized spline* menggunakan *Bayesian Information Criterion* (BIC) minimum diperoleh kombinasi orde dan jumlah knot optimal dengan orde berjumlah 2 sedangkan knot berjumlah 1 yang diperoleh dari nilai BIC minimum sebesar 543.3806, ketika  $\lambda$  setara dengan 0,001. Berdasarkan uji kelayakan model menggunakan nilai MAPE, kinerja model regresi *penalized spline* dalam meramalkan data Pengeluaran per Kapita terhadap persentase IPM terbukti efektif.

3. Dari proses analisis, didapatkan informasi bahwa ketika nilai Pengeluaran per Kapita yang Disesuaikan terhadap persentase Indeks Pembangunan Manusia (IPM) mencapai nilai minimum sebesar 8173.879, maka Pengeluaran per Kapita akan mengalami peningkatan sebesar 8173.879. Selain itu, saat nilai pengeluaran per kapita terhadap persentase IPM mencapai nilai maksimum yakni 24221, dapat diinterpretasikan bahwa dengan menggantikan nilai maksimum maka Pengeluaran per Kapita akan mengalami peningkatan sebesar 24217.31. Berikut model regresi *penalized spline* yang memiliki hasil optimal yaitu  $\hat{y} = 8173.879Z_{-1,2}(x) + 8823.742Z_{0,2}(x) + 24217.31Z_{-1,2}(x)$ .

## 5.2. Saran

1. Penelitian ini memiliki batasan jumlah variabel independen sebanyak satu, untuk penelitian selanjutnya dapat menggunakan lebih dari satu variabel independen.
2. Jumlah orde yang digunakan pada penelitian ini sebanyak 2 sampai 4 dan jumlah knot sebanyak 1 sampai 3 dengan batas parameter penghalus dari 0.001 hingga 0.99. Untuk peneliti selanjutnya bisa mengembangkan jumlah orde, knot, dan parameter penghalus yang digunakan.
3. Banyak metode yang dapat digunakan selain metode *Bayesian Information Criterion* (BIC).

## DAFTAR PUSTAKA

- Agustina, N., Suparti, S., & Mukid, M. A. (2015). Pemodelan data indeks harga saham gabungan menggunakan regresi penalized spline. *Jurnal Gaussian*, 4(3):603–612.
- Asikin, M. I., Bachrudin, A., & Tantular, B. (2020). Pemodelan balita stunting di jawa barat menggunakan generalized additive models penalized splines. In *Prosiding Seminar Nasional Statistika Departemen Statistika FMIPA Universitas Padjadjaran*, volume 9, pages 26–26.
- Barry, P. (1996). de boor-fix dual functionals and algorithms for tchebycheffian b-spline curves. *Constructive Approximation*, 12:385–408.
- Berry, L. N. & Helwig, N. E. (2021). Cross-validation, information theory, or maximum likelihood? a comparison of tuning methods for penalized splines. *Stats*, 4(3):701–724.
- BPS (2022). *Indeks Pembangunan Manusia 2022*. BPS.
- Budiantara, I. N., Suryadi, F., Otok, B. W., & Guritno, S. (2006). Pemodelan b-spline dan mars pada nilai ujian masuk terhadap ipk mahasiswa jurusan disain komunikasi visual uk. petra surabaya. *Jurnal Teknik Industri*, 8(1):1–13.
- Durbán, M., Currie, I., & Eilers, P. (2002). Using p-splines to smooth two-dimensional poisson data. In *Proceedings of 17th International Workshop on Statistical Modelling, Chania, Crete*, pages 207–214.
- Eilers, P. H. & Marx, B. D. (1996). Flexible smoothing with b-splines and penalties. *Statistical science*, 11(2):89–121.
- Eilers, P. H. & Marx, B. D. (2010). Splines, knots, and penalties. *Wiley interdisciplinary reviews: Computational statistics*, 2(6):637–653.

- Eubank, R. L. (1999). *Nonparametric regression and spline smoothing*. CRC press.
- Kurniasari, W., Kusnandar, D., & Sulistianingsih, E. (2019). Estimasi parameter regresi spline dengan metode penalized spline. *Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya*, 8(2).
- Lewis, C. D. (1982). *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting. (No Title)*.
- Qudratullah, M. F. (2013). Analisis regresi terapan: Teori, contoh kasus, dan aplikasi dengan spss. *Yogyakarta: Andi*.
- Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric regression*. Number 12. Cambridge university press.
- Santoso, S. (2019). *Mahir statistik parametrik*. Elex Media Komputindo.
- Sugiarti, H. (2012). Kesesuaian metode regresi nonparametrik spline, b-spline, dan p-spline dalam menduga kurva regresi. *Pusat Penelitian dan Pengabdian Kepada Masyarakat, Universitas Terbuka*.
- Supandi, E. D. (2020). *Statistika dan Terapannya*. Refika Aditama.
- Zia, N. G., Suparti, S., & Safitri, D. (2017). Pemodelan regresi spline menggunakan metode penalized spline pada data longitudinal (studi kasus: Harga penutupan saham Iq45 sektor keuangan dengan kurs usd terhadap rupiah periode januari 2011-januari 2016). *Jurnal Gaussian*, 6(2):221–230.