

SKRIPSI

**OPTIMASI MODEL REGRESI *PENALIZED SPLINE*
MENGUNAKAN METODE *AKAIKE INFORMATION*
CRITERION (AIC) PADA DATA *CROSS SECTION***

(STUDI KASUS : TINGKAT PARTISIPASI ANGKATAN KERJA (TPAK) TERHADAP 3
PROVINSI DENGAN PERSENTASE TINGKAT KEMISKINAN TERTINGGI DI INDONESIA

TAHUN 2021)



ARA SYAIFHA NAFIA
NIM. 19106010028
STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

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

2023

**OPTIMASI MODEL REGRESI *PENALIZED SPLINE*
MENGUNAKAN METODE *AKAIKE INFORMATION*
CRITERION (AIC) PADA DATA *CROSS SECTION***

(STUDI KASUS : TINGKAT PARTISIPASI ANGKATAN KERJA (TPAK) TERHADAP 3
PROVINSI DENGAN PERSENTASE TINGKAT KEMISKINAN TERTINGGI DI INDONESIA
TAHUN 2021)

Skripsi

Diajukan untuk Memenuhi Salah Satu Syarat Memperoleh Derajat Sarjana

Matematika

Program Studi Matematika



ARA SYAIFHA NAFIA

NIM. 19106010028

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 : Ara Syaifha Nafia
NIM : 19106010028
Judul Skripsi : Optimasi Model Regresi *Penalized Spline* Menggunakan Metode *Akaike Information Criterion (AIC)* Pada Data *Cross section*

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

Wassalamu'alaikum wr. wb.

Yogyakarta, 09 Agustus 2023

Pembimbing I

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-2213/Un.02/DST/PP.00.9/08/2023

Tugas Akhir dengan judul : OPTIMASI MODEL REGRESI PENALIZED SPLINE MENGGUNAKAN METODE AKAIKE INFORMATION CRITERION (AIC) PADA DATA CROSS SECTION (STUDI KASUS : TINGKAT PARTISIPASI ANGKATAN KERJA (TPAK) TERHADAP 3 PROVINSI DENGAN PERSENTASE TINGKAT KEMISKINAN TERTINGGI DI INDONESIA TAHUN 2021)

yang dipersiapkan dan disusun oleh:

Nama : ARA SYAIFHA NAFIA
Nomor Induk Mahasiswa : 19106010028
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: 64e32938261f8



Penguji I

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

Valid ID: 64e2d046a7a21



Penguji II

Pipit Pratiwi Rahayu, S.Si., M.Sc.
SIGNED

Valid ID: 64e2ea954b40c



Yogyakarta, 15 Agustus 2023

UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

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

Valid ID: 64e4159d61787

SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Ara Syaifha Nafia

NIM : 19106010028

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



Ara Syaifha Nafia

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PERSEMBAHAN

*Tiada lembar yang paling indah dalam karya tulis saya
kecuali lembar persembahan ini sebagai tanda bukti
cinta dan kasih kepada :*

*Kedua orang tua, bapak Alm. Syaiful Barry ; Dwi Yanto
dan ibu Halilintar Meganingrum, yang telah memberikan
do'a dan dukungan yang tak henti.*

*Keluarga besar, sahabat, teman, dan setiap orang yang
pernah saya kenal.*

*Almamater Universitas Islam Negeri Sunan Kalijaga
Yogyakarta.*

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

MOTTO

"Allah tidak akan pernah membebani seseorang melainkan sesuai dengan kemampuannya."

(QS. Al-Baqarah: 286)

SKRIPSI YANG BAIK ITU SKRIPSI YANG SELESAI

You don't have to be perfect, you just need to be "you" in the perfect version



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PRAKATA

Assalamu'alaikum Wr.Wb

Alhamdulillah, puji syukur kepada Allah SWT yang telah memberikan rahmat dan karunianya berupa keimanan, kekuatan, kesabaran, kelancaran dan keselamatan sehingga penulis dapat menyelesaikan skripsi yang berjudul “Optimisasi Model Regresi Penalized Spline (P-Spline) Menggunakan Model *Akaike Information Criterion* (AIC)”. Skripsi ini ditulis guna memenuhi syarat memperoleh derajat kesarjanaan ilmu matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta. Penulis menyadari bahwa selama penyusunan skripsi ini telah banyak mendapatkan bimbingan dan motivasi dari berbagai pihak. Untuk itu pada kesempatan ini penulis bermaksud menyampaikan ucapan terima kasih kepada :

1. Ibu Prof. Dr. Dra. Hj. Khurul Wardati, M.Si., selaku Dekan Fakultas Sains dan Teknologi, UIN Sunan Kalijaga Yogyakarta.
2. Bapak Muchammad Abrori, S.Si., M.Kom., selaku Ketua Program Studi Matematika beserta jajarannya.
3. Ibu Sri Utami Zuliana, S.Si., M.Sc., Ph.D., selaku Dosen Pembimbing skripsi yang telah berkenan memberikan bimbingan, motivasi, arahan, waktu, pikiran, dan kesabaran yang besar bagi penulis, sehingga dapat menyelesaikan skripsi ini dengan baik.
4. Ibu Malahayati, S.Si., M.Sc., selaku Dosen Pembimbing Akademik yang selalu memberikan arahan dan nasihat kepada penulis untuk menyelesaikan skripsi dengan baik.
5. Seluruh dosen dan staf Fakultas Sains dan Teknologi yang selalu memberikan ilmu dan pelayanan yang baik selama perkuliahan sampai pada selesainya

penyusunan skripsi ini.

6. Kedua orang tua, papa Alm. Syaiful Barry dan Dwi Yanto ; mama Halilintar yang selalu memberikan cinta dan kasih sayang, selalu mendukung dan mendoakan yang terbaik untuk saya.
7. Tante saya, Dr.Insanul Qisti Barriyah, M.Sn. dan keluarga Barry lainnya yang selalu menjadi *reminder* bagi penulis dan memberikan kasih sayang yang besar seperti anaknya sendiri.
8. Teman-teman satu bimbingan tugas akhir (bimbingan Ibu Zuliana) Devita, Indriani, Tsalisa, Putri, Shelia dan Ibnu yang selalu bersama berbagi ilmu.
9. Sahabat saya Pulu - Pulu (Devita, Indri, Putri, dan Tsalisa) dan sahabat UKM Kalimasada.
10. Teman-teman Matematika angkatan 2019 yang telah kebersamai selama 4 tahun terakhir dan telah memberikan kesan terbaik, selalu membagi ilmu dan pengalaman baik dalam dunia perkuliahan maupun di luar perkuliahan.
11. Semua pihak yang tidak dapat penulis sebutkan satu-persatu, terima kasih.

Semoga skripsi ini dapat memberikan manfaat bagi penulis khususnya, dan bagi semua kalangan yang membutuhkan umumnya. Akhir kata, semoga Allah SWT selalu melindungi dan melimpahkan rahmat serta hidayah-Nya kepada kita semua. *Aamiin ya Allah ya robbal alamiin.*

Wassalamu'alaikum Wr. Wb.

Yogyakarta, 07 Agustus 2023

Penulis

DAFTAR ISI

| | |
|--|-------------|
| HALAMAN JUDUL | i |
| SURAT PERSETUJUAN | ii |
| SURAT PENGESAHAN | iii |
| SURAT PERNYATAAN KEASLIAN | iv |
| HALAMAN PERSEMBAHAN | v |
| HALAMAN MOTTO | vi |
| PRAKATA | vii |
| DAFTAR ISI | ix |
| DAFTAR TABEL | xii |
| DAFTAR GAMBAR | xiii |
| DAFTAR LAMBANG | xiv |
| INTISARI | xv |
| ABSTRACT | xvi |
| I PENDAHULUAN | 1 |
| 1.1. Latar Belakang Masalah | 1 |
| 1.2. Batasan Masalah | 4 |
| 1.3. Rumusan Masalah | 4 |
| 1.4. Tujuan Penelitian | 5 |
| 1.5. Manfaat Penelitian | 5 |
| 1.6. Tinjauan Pustaka | 5 |
| 1.7. Sistematika Penulisan | 6 |
| II DASAR TEORI | 8 |
| 2.1. <i>Cross Section</i> | 8 |

| | |
|---|-----------|
| 2.2. Analisis Regresi | 8 |
| 2.3. Analisis Regresi Linier Sederhana | 10 |
| 2.4. Estimasi Regresi Linier | 10 |
| 2.4.1. (<i>Ordinary Least Square (OLS)</i>) | 11 |
| 2.4.2. <i>Maximum Likelihood</i> | 11 |
| 2.5. Analisis Regresi Parametrik | 12 |
| 2.5.1. Uji Normalitas | 13 |
| 2.5.2. Uji Autokorelasi | 13 |
| 2.5.3. Uji Heterokedastisitas | 14 |
| 2.5.4. Uji Multikolinieritas | 14 |
| 2.6. Analisis Regresi Nonparametrik | 14 |
| 2.7. Regresi <i>Spline</i> | 15 |
| 2.8. Regresi B-Spline | 16 |
| 2.9. Basis-Basis <i>B-Spline</i> | 18 |
| 2.10. Regresi <i>Penalized Spline</i> | 18 |
| 2.11. Optimisasi Parameter Penghalus (λ) pada <i>Penalized Spline</i> | 20 |
| 2.11.1. Pengujian Serentak | 21 |
| 2.12. <i>Mean Absolute Percentage Error (MAPE)</i> | 22 |
| 2.13. Indeks Kemiskinan menurut Badan Pusat Statistik | 23 |
| 2.14. Faktor-Faktor Mempengaruhi Kemiskinan | 24 |
| III METODE PENELITIAN | 25 |
| 3.1. Data dan Jenis Penelitian | 25 |
| 3.2. Metode Pengumpulan Data | 26 |
| 3.3. Sumber Data dan Variabel Penelitian | 26 |
| 3.4. Pengolahan Data | 26 |
| 3.5. <i>Software</i> yang digunakan | 27 |

| | |
|---|-----------|
| 3.6. Source Code | 27 |
| IV HASIL DAN PEMBAHASAN | 28 |
| 4.1. Deskripsi Data | 28 |
| 4.2. Uji Asumsi Klasik | 30 |
| 4.3. Pemilihan Orde, Jumlah Knot, dan Lamda Optimal | 30 |
| 4.3.1. Interpretasi Model | 32 |
| 4.4. Pengujian Signifikansi Parameter | 35 |
| 4.5. Pengujian Kelayakan Model | 35 |
| V PENUTUP | 38 |
| 5.1. Kesimpulan | 38 |
| 5.2. Saran | 39 |
| LAMPIRAN A | 42 |
| LAMPIRAN B | 43 |
| LAMPIRAN C | 57 |
| Curriculum Vitae | 58 |

DAFTAR TABEL

| | | |
|-----|--|----|
| 2.1 | Analisis Variansi (ANOVA) Signifikansi Parameter | 22 |
| 4.1 | Data Kemiskinan di Provinsi Papua, Papua Barat, dan NTT tahun 2021 | 28 |
| 4.2 | Titik Knot Optimal untuk Masing-Masing Orde | 31 |
| 4.3 | Estimasi Parameter Menggunakan Metode AIC | 32 |
| 4.4 | Analisis Uji Serentak Parameter Model <i>P-Spline</i> | 35 |
| 4.5 | Uji Kinerja Model Menggunakan Nilai <i>Mean Percentage Error</i> (MA- PE) | 36 |



DAFTAR GAMBAR

- 4.1 Matrix Korelasi yang Mempengaruhi Kemiskinan 29
- 4.2 Grafik hubungan antara lambda dengan nilai AIC 31
- 4.3 Prediksi Persentase Kemiskinan Provinsi Jawa Tengah tahun 2021 . 36



DAFTAR LAMBANG

| | | |
|---------------|---|------------------------------|
| ϵ | = | Nilai <i>error</i> |
| x | = | Variable Prediktor |
| y | = | Variable Respon |
| $f(x)$ | = | Fungsi Regresi |
| β_0 | = | Konstanta |
| β_i | = | Koefisien Pada Variable Ke-1 |
| K | = | Titik Knot |
| v | = | Letak Titik Knot |
| K_v | = | Titik Knot ke- v |
| m | = | Orde |
| k | = | Banyak Titik Knot |
| λ | = | Parameter Penghalus |
| \hat{y} | = | Estimasi Variable Prediktor |
| $\hat{\beta}$ | = | Parameter |
| OLS | = | Ordinary Least Square |
| B | = | Basis <i>B-spline</i> |
| Δ_d | = | <i>Different Operator</i> |
| D | = | Matriks Diagonal |
| L | = | Fungsi <i>Likelihood</i> |
| I | = | Matriks Identitas |

INTISARI

OPTIMASI MODEL REGRESI *PENALIZED SPLINE* MENGGUNAKAN METODE *AKAIKE INFORMATION CRITERION (AIC)* PADA DATA *CROSS SECTION*

(STUDI KASUS : TINGKAT PARTISIPASI ANGKATAN KERJA (TPAK) TERHADAP 3
PROVINSI DENGAN PERSENTASE TINGKAT KEMISKINAN TERTINGGI DI INDONESIA
TAHUN 2021)

Oleh

ARA SYAIFHA NAFIA

NIM. 19106010028

Penalized Spline yaitu penggabungan regresi dengan basis *B-Spline* untuk kemulusan kurva. Tujuan penelitian ini adalah untuk mendapatkan model persentase kemiskinan dengan variabel TPAK menggunakan regresi *Penalized Spline*. Model regresi *Penalized Spline* optimal diperoleh berdasarkan pemilihan kombinasi orde, jumlah titik knot, dan lambda yang optimal menggunakan optimisasi *Akaike Information Criterion (AIC)*. Selanjutnya diperoleh estimasi parameter model dan dilakukan pengujian kelayakan model untuk melihat kinerja model menggunakan nilai *Mean Absolute Percentage Error (MAPE)*. Diperoleh kombinasi orde dan jumlah titik knot optimal terdapat pada orde 2 (linier), 2 titik knot, dan 0.001 lambda dengan nilai AIC sebesar 435.4894 dan nilai MAPE sebesar 14.48345. Model optimal yang diperoleh adalah: $\hat{y} = 17.69902B_{-1,2}(x) + 24.35812B_{0,2}(x) + 18.76323B_{1,2}(x) + 40.36929B_{2,2}(x)$

Kata Kunci: AIC, *Cross Section*, MAPE, Regresi *Penalized Spline*

ABSTRACT

**OPTIMIZATION OF THE PENALIZED SPLINE REGRESSION MODEL
USING THE AKAIKE INFORMATION CRITERION (AIC) METHOD
(CASE STUDY: LABOR FORCE PARTICIPATION RATE (TPAK) IN 3
PROVINCES WITH THE HIGHEST PERCENTAGE OF POVERTY LEVELS IN
INDONESIA, 2021)**

By

ARA SYAIFHA NAFIA

NIM. 19106010028

Penalized Spline is a combination of regression with a B-Spline basis for a smooth curve. The purpose of this study was to obtain a poverty percentage model with the TPAK variable using the Penalized Spline regression. The optimal Penalized Spline regression model obtained by selecting of the optimal combination of order, amount of knot points, and using Akaike Information Criterion (AIC) optimization to get the lambda optimal. After estimation of model parameters was obtain and then testing the model feasibility testing to carried out the model performance by using the Mean Absolute Percentage Error (MAPE) value. The optimal combination of order and amount of knot points obtained at order 2 (linear), 2 knot points, and 0.001 lambda with an AIC value of 435.4894 and a MAPE value of 14.48345. The optimal model obtained : $\hat{y} = 17.69902B_{-1,2}(x) + 24.35812B_{0,2}(x) + 18.76323B_{1,2}(x) + 40.36929B_{2,2}(x)$

Keyword: *AIC, Cross Section, MAPE, Penalized Spline*

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Cross Section merupakan sekumpulan data yang terdiri dari banyak objek atau variabel yang diukur berdasarkan satu waktu yang sama. Analisis *Cross Sectional* merupakan studi yang mengkaji mengenai dinamika dan hubungan antara faktor risiko dengan menggunakan metode observasi atau pengumpulan data. Beberapa petinggi perusahaan atau instansi membutuhkan analisis tersebut guna melihat suatu perkembangan terhadap industrinya. Pemodelan tipe data ini salah satunya dengan menggunakan analisis regresi yaitu dengan mengetahui pola dan perkiraan nilai korelasi antara variabel respon dan variabel prediktor.

Pada umumnya analisis regresi terbagi menjadi 3 yaitu analisis regresi parametrik, analisis semiparametrik, dan analisis regresi nonparametrik. Analisis regresi semiparametrik merupakan gabungan dari analisis regresi parametrik dan nonparametrik. Analisis regresi parametrik merupakan analisis yang paling umum digunakan jika bentuk fungsi $f(x)$ diketahui dari informasi sebelumnya berdasarkan teori masa lampau. Dalam regresi untuk mendeskripsikan kedua jenis variabel digunakan sebuah kurva regresi. Pola hubungan yang terasumsikan pada kurva regresi membentuk linier, kuadratik, kubik, dan lain-lain. Dalam mendapatkan model regresi yang terbaik pada analisis regresi parametrik harus memenuhi beberapa asumsi klasik yaitu uji normalitas, uji autokorelasi, uji heterokedastisitas, dan uji multikorelasi. Pelanggaran asumsi mengakibatkan model regresi yang dihasilkan kurang baik, sehingga apabila salah satu asumsi tidak terpenuhi, maka menggunakan analisis regresi nonparametrik (Qudratullah, 2013). Estimasi analisis regresi nonparametrik dilakukan melalui data pengamatan dengan beberapa metode dalam pembentukan model yaitu spline, kernel, MARS, deret fourier, wavelets, dan lain-lain.

Regresi *Spline* merupakan suatu pendekatan kearah plot data dengan mengestimasi kemulusan kurva, dimana kurva tersebut terbentuk dari penggabungan titik knot. Dalam pembentukan kurva regresi *spline* memberikan fleksibilitas yang lebih besar, sifat tersebutlah memungkinkan model regresi menyesuaikan diri secara efektif terhadap karakteristik data. Regresi *spline* memiliki basis umum yaitu basis fungsi *truncated* dan basis fungsi *B-Spline*. Adapun kekurangan dari basis *truncated* yaitu apabila menggunakan orde tinggi, jumlah knot akan bertambah dan letak knot semakin berdekatan sehingga akan menciptakan matriks yang hampir singular yang nantinya membuat persamaan yang tidak stabil. Untuk mengatasi permasalahan tersebut maka disarankan menggunakan basis *B-Spline* karena mampu membentuk persamaan matriks yang hampir singular (Eubank, 1999).

Pada regresi *B-Spline* dianggap sensitif terhadap knot. Apabila diberi knot terlalu banyak akan menghasilkan model *overfit* dan terlalu kecil akan menghasilkan model *underfit*. Dikarenakan hasil dari kurva regresi *B-Spline* masih terlihat kasar, sehingga ditambahkan kemulusan *penalty* dengan menggunakan regresi *Penalized Spline* (Eilers & Marx, 1996). Pemodelan menggunakan regresi *Penalized Spline* bertumpu pada kombinasi orde, jumlah knot yang optimal, dan besarnya lambda. Adapun dalam pemilihan jumlah knot yang optimal untuk regresi *Penalized spline* yaitu dengan menggunakan beberapa metode dengan melihat nilai minimumnya, beberapa optimisasi yang digunakan, seperti *Akaike Information Criterion (AIC)*. Penulis menggunakan optimasi *AIC* dengan kombinasi orde 2 hingga 4 dan jumlah titik knot 1 hingga 4 titik knot. Pemilihan orde dan titik knot terbatas dikarenakan mempersingkat waktu dalam pengolahan data. Selain itu, kurva dapat mengalami *overfitting* apabila pemilihan titik knot terlalu banyak.

Penerapan *Penalized Spline* sudah banyak digunakan untuk penelitian pada permasalahan sehari-hari misalnya Indeks Pembangunan Manusia (IPM), tekanan darah, harga penutupan saham, dan tingkat kemiskinan. Kemiskinan merupakan rendahnya pendapatan dalam memenuhi kebutuhan dasar (pangan, sandang, dan papan), rendahnya sarana pelayanan kesehatan maupun pendidikan, serta kapasitas dasar masyarakat secara memadai. Permasalahan kemiskinan terjadi tidak hanya di

lingkup nasional saja melainkan sudah menjadi permasalahan dunia. Perserikatan Bangsa-Bangsa (PBB) pada September 2000 melaporkan mengenai kebijakan MDGs (*Millenium Development Goals*) yang memberikan pernyataan bahwa sasaran utama dari MDGs adalah membasmi kemiskinan dan kelaparan ekstrem (Nations, 2007).

Indonesia masuk dalam kategori 100 negara termiskin didunia. Indonesia merupakan negara yang berkembang dimana menjadi salah satu negara terbelakang yang masih bertopang pada negara maju. Permasalahan di Indonesia identik dengan kemiskinan yang akan menjadi perhatian serius dari pemerintah. Jumlah penduduk yang banyak dengan sebagian besar penduduknya memiliki tingkat ketenagakerjaan yang rendah akan memicu adanya kesenjangan sosial dan terjadi kemiskinan. Indonesia bagian timur merupakan suatu wilayah yang berada di bagian timur Indonesia, wilayah ini dikategorikan memiliki persentase kemiskinan tertinggi di Indonesia. Berdasarkan pendataan Badan Pusat Statistika (BPS) pada September 2021 provinsi Papua dengan menduduki persentase penduduk miskin tertinggi sebesar 26,86%, Papua Barat menduduki tingkat tertinggi kedua sebesar 21,84%, dan Nusa Tenggara Timur menduduki tingkat tertinggi ketiga sebesar 20,44%.

Ketangakerjaan dan kemiskinan menjadi suatu permasalahan yang kompleks di negara Indonesia terutama di wilayah timur. Salah satu indikator yang menentukan tingkat kemakmuran sosial ekonomi suatu rumah tangga adalah pengasilan utama. Karakteristik ketenagakerjaan menjelaskan adanya ketidaksamaan antara rumah tangga miskin atau tidak miskin terlihat dari lapangan usaha atau sumber penghasilan utama rumah tangga dan status pekerjaan (BPS, 2021). Sumber penghasilan mencapai maksimum apabila diimbangi dengan penggunaan tenaga kerja penuh dapat diwujudkan. Tingkat Partisipasi Angkatan Kerja (TPAK) merupakan perbandingan antara angkatan kerja dengan penduduk usia kerja (BPS, 2000). Angkatan kerja merupakan penduduk yang berkeja dan penduduk yang mencari pekerjaan sedangkan penduduk usia kerja merupakan penduduk yang berusia lebih dari 15 tahun. Semakin tinggi TPAK pada suatu daerah, semakin tinggi pula jumlah tenaga kerja yang tersedia untuk melakukan produksi barang dan jasa dalam jalanya

perekonomian suatu daerah, begitu pula sebaliknya (BPS, 2020). Oleh karena itu, TPAK menjadi salah satu indikasi yang berpengaruh pada tingkat kemiskinan.

Berdasarkan uraian di atas, maka penulis akan membahas persentase kemiskinan di Indonesia bagian timur menggunakan regresi nonparametrik *Penalized Spline*. Studi ini mengkaji tentang persentase kemiskinan di Provinsi Papua, Papua Barat, dan Nusa Tenggara Timur sebagai variabel respon, sedangkan Tingkat Partisipasi Angkatan Kerja (TPAK) sebagai variabel prediktor, maka dalam penelitian ini mengambil judul “*Optimasi Model Regresi Penalized Spline (P-Spline) Dengan Metode Akaike Information Criterion (AIC) Pada Data Cross Section*”

1.2. Batasan Masalah

Penulisan tugas akhir ini perlu ditentukan batasan masalah, pembahasan difokuskan pada pokok-pokok bahasan sehingga tidak menyimpang dari tujuan semula. Selanjutnya batasan masalah pada tugas akhir ini sebagai berikut:

1. Pemilihan model optimasi regresi *Penalized Spline* menggunakan metode *Akaike Information Criterion (AIC)*
2. Orde yang digunakan orde 2 hingga 4, jumlah knot 1 hingga 4, dan batas lambda 0.001 hingga 0.99.
3. Software yang digunakan adalah R Studio versi 2022.12.0+353 dan bahasa pemrograman R versi 4.2.2

1.3. Rumusan Masalah

Berdasarkan latar belakang yang ada maka penulis merumuskan masalah yang akan dibahas pada penelitian ini, sebagai berikut :

1. Bagaimana langkah-langkah analisis regresi nonparametrik *Penalized Spline*?
2. Bagaimana memodelkan persentase kemiskinan di Provinsi Papua, Papua Barat, dan NTT dengan pendekatan regresi nonparametrik *Penalized Spline* dengan metode *Akaike Information Criterion (AIC)* ?

1.4. Tujuan Penelitian

Adapun tujuan penelitian ini sebagai berikut:

1. Mengkaji dan memahami tentang model regresi *Penalized Spline*
2. Mengkaji optimasi *Akaike Information Criterion* (AIC) dalam regresi nonparametik *Penalized Spline*
3. Mengaplikasikan model *Penalized Spline* pada data persentase kemiskinan di Provinsi Papua, Papua Barat, dan NTT tahun 2021.

1.5. Manfaat Penelitian

Adapun manfaat penelitian ini sebagai berikut :

1. Menambah pengetahuan mengenai regresi *Penalized Spline*, optimisasinya, dan penerapan teori dalam kasus Tingkat Partisipasi Angkatan Kerja (TPAK) terhadap presentase tingkat kemiskinan Indonesia.
2. Penelitian ini dapat digunakan sebagai referensi dan pengembangan lebih lanjut bagi pembaca.

1.6. Tinjauan Pustaka

Tinjauan pustaka pada penelitian ini didiskripsikan an telah melalui buku, makalah, skripsi, jurnal matematika, maupun sumber lainnya yang terdapat relevansinya dengan objek pembahasan. Adapun tinjauan pustaka ini antara lain:

1. (Siregar & Jatmiko, 2019) membahas penelitian mengenai hubungan Indeks Pembangunan Manusia (IPM) dengan persentase penduduk miskin kabupaten/kota di Indonesia menggunakan perbandingan regresi *B-Spline* dan regresi *P-Spline* dengan metode yang digunakan sebagai parameter penghalus yang optimal diperoleh berdasarkan nilai *Generalized Cross Validation* (GCV) dan RSS minimum. Adapun pemilihan metode terbaik untuk mengetahui seberapa akurat peramalan yang dihasilkan menggunakan MSE.

2. (Zia et al., 2017) membahas penelitian mengenai harga penutupan saham LQ45 sektor keuangan dengan kurs USD terhadap rupiah, menggunakan pendekatan regresi *Penalized Spline* (P-Spline) dengan metode yang digunakan sebagai parameter penghalus yang optimal diperoleh berdasarkan nilai *Generalized Cross Validation* (GCV) minimum. Adapun pemilihan metode terbaik untuk mengetahui seberapa akurat peramalan yang dihasilkan menggunakan R^2 dan MAPE.
3. (Kurniasari et al., 2019) membahas penelitian mengenai Indeks Pembangunan Manusia (IPM) menurut provinsi di Indonesia tahun 2015, menggunakan estimasi parameter regresi *Spline* dengan metode *Penalized Spline*. Adapun pemilihan model terbaik pada setiap variabel dengan nilai *Generalized Cross Validation* (GCV) minimum untuk mengetahui seberapa akurat peramalan yang dihasilkan menggunakan MAPE.
4. (Sriliiana et al., 2017) membahas penelitian mengenai pemetaan kemiskinan di Kabupaten Mukomuko, menggunakan metode *small area estimation* (SAE) dengan pendekatan regresi *Penalized Spline* (P-spline). Penentuan jumlah dan keberadaan titik knot dengan menggunakan *Generalized Cross Validation*(GCV). Adapun pemilihan kriteria model terbaik dengan menggunakan *Akaike Information Criterion* (AIC).

1.7. Sistematika Penulisan

Dalam mempermudah pemahaman penulisan penelitian ini dengan cara yang secara komprehensif, demikian penulis menggambarkan sistem penulisan sebagai berikut :

- BAB 1** : Pada bab ini membahas tentang latar belakang masalah, batasan masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, tinjauan pustaka, metode penelitian, dan sistematika penulisan.
- BAB 2** : Pada bab ini membahas tentang teori-teori pendukung terhadap penelitian ini.
- BAB 3** : Pada bab ini akan dibahas tentang Data dan Jenis Penelitian, Metode Pengumpulan Data, Populasi dan Sampel, Variabel Penelitian, dan Pengolahan Data.
- BAB 4** : Pada bab ini membahas hasil penelitian dari pembahasan permasalahan.
- BAB 5** : Pada bab ini berisikan tentang kesimpulan penelitian dan saran dari penulis terhadap pengembangan penelitian.

BAB V

PENUTUP

5.1. Kesimpulan

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

1. Dalam memodelkan data menggunakan regresi *P-Spline*, beberapa langkah analisis yang dilakukan dalam penelitian ini yaitu menentukan perkiraan model regresi *P-Spline* berdasarkan orde dan jumlah titik knot yang diinginkan, menentukan kombinasi orde dan jumlah titik knot optimal berdasarkan nilai AIC minimum, mengestimasi parameter model regresi *P-Spline* terbaik berdasarkan kriteria AIC, menguji signifikansi parameter terhadap model yang dibangun menggunakan uji serentak, dan menguji kelayakan model berdasarkan nilai MAPE. Tahapan analisis tersebut menghasilkan informasi yaitu ketika nilai TPAK terhadap persentase kemiskinan bernilai minimum yaitu 56.39%, maka nilai minimum tersebut dapat disubstitusikan ke dalam persamaan (4.12) sehingga dapat diketahui bahwa TPAK akan meningkat sebesar 17.70% . Selain itu, ketika nilai TPAK terhadap persentase kemiskinan bernilai maksimum yaitu 97.93% , maka nilai maksimum tersebut dapat disubstitusikan ke dalam persamaan (4.14) sehingga dapat diketahui bahwa TPAK akan meningkat sebesar 40.37%.
2. Kombinasi orde dan jumlah titik knot yang optimal diperoleh dengan melihat nilai AIC minimum pada setiap orde dan jumlah titik knot yang diuji. Diperoleh kombinasi orde dan jumlah titik knot optimal terdapat pada orde 2 dan jumlah titik knot sebanyak 2 dengan nilai AIC sebesar 435.4894.
3. Model regresi *P-Spline* dengan diperoleh nilai MAPE sebesar 14.682 memiliki kinerja yang baik dalam melakukan peramalan pada data TPAK terhadap

persentase kemiskinan. Berikut model regresi *P-Spline* optimal.

$$\hat{y} = 17.69902B_{-1,2}(x) + 24.35812B_{0,2}(x) + 18.76323B_{1,2}(x) + 40.36929B_{2,2}(x)$$

5.2. Saran

1. Penelitian ini menggunakan satu variabel yaitu data TPAK terhadap persentase kemiskinan Papua, Papua Barat, dan NTT yang terbatas pada kombinasi orde 2 sampai orde 4 dengan 1 sampai 4 jumlah titik knot. Penulis berharap penelitian yang telah dilakukan ini dapat menjadi bahan rujukan untuk pengembangan penelitian selanjutnya, baik dari segi penambahan kombinasi orde dan knot maupun variabel yang digunakan.
2. Metode yang digunakan dalam menentukan model regresi P-Spline terbaik adalah menggunakan optimisasi Akaike Information Criterion (AIC). Optimisasi ini merupakan satu dari banyaknya pilihan optimisasi yang dapat digunakan. Penulis berharap adanya variasi optimisasi yang digunakan untuk membentuk model regresi B-Spline terbaik dalam pengembangan penelitian selanjutnya.

DAFTAR PUSTAKA

- BPS, B. P. S. (2000). Sensus penduduk 2000.
- BPS, B. P. S. (2020). Provinsi papua barat dalam angka 2020.
- BPS, B. P. S. (2021). Perhitungan dan analisis kemiskinan makro indonesia 2021.
- Burnham, K. P. & Anderson, D. R. (2004). Multimodel inference: understanding aic and bic in model selection. *Sociological methods & research*, 33(2):261–304.
- De Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean absolute percentage error for regression models. *Neurocomputing*, 192:38–48.
- Eilers, P. H. & Marx, B. D. (1996). Flexible smoothing with b-splines and penalties. *Statistical science*, 11(2):89–121.
- Eubank, R. L. (1999). *Nonparametric regression and spline smoothing*. CRC press.
- Hasan, M. I. (2002). Pokok-pokok materi metodologi penelitian dan aplikasinya.
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer.
- 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*. Butterworth-Heinemann.
- Nations, U. (2007). *Governance for the Millennium Development Goals: Core Issues and Good Practices*. New York, United Nations Publication.

- Quadratullah (2013). *Analisis Regresi Terapan*. Yogyakarta: SUKA-Press UIN Sunan Kalijaga.
- Rosadi, D. (2011). *Analisis ekonometrika runtun waktu terapan dengan R : aplikasi untuk bidang ekonomi, bisnis, dan keuangan*. Yogyakarta.
- Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric regression*. Number 12. Cambridge university press.
- Siregar, R. S. K. & Jatmiko, Y. A. (2019). Perbandingan regresi b-splines dan p-splines pada hubungan indeks pembangunan manusia dan persentase penduduk miskin kabupaten/kota di indonesia: Comparison of b-splines and p-splines regression on the relationship between human development index and the percentage of poor districts/cities in indonesia. *Emasains: Jurnal Edukasi Matematika dan Sains*, 8(1):101–111.
- Sriliiana, I., Sunandi, E., & Rafflesia, U. (2017). Pemodelan kemiskinan di provinsi bengkulu menggunakan small area estimation dengan pendekatan semiparametrik penalized spline. *Indonesian Journal of Mathematics and Natural Sciences*, 40(2):134–140.
- SugiyonoJ (2016). *Metode Penelitian Kuantitatif*. Number 225. Bandung : IKAPI.
- 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.