



TUGAS AKHIR - KS184822

**PEMODELAN PERSENTASE BALITA GIZI BURUK
DI INDONESIA TAHUN 2018 MENGGUNAKAN
REGRESI SEMIPARAMETRIK *SPLINE TRUNCATED***

**WANINDYATAMI FIRSTIDI PUTRI
NRP 062115 4000 0037**

**Dosen Pembimbing
Jerry Dwi Trijoyo Purnomo, S.Si., M.Si., Ph.D.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS SAINS DAN ANALITIKA DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2020**



TUGAS AKHIR - KS184822

**PEMODELAN PERSENTASE BALITA GIZI BURUK
DI INDONESIA TAHUN 2018 MENGGUNAKAN
REGRESI SEMIPARAMETRIK *SPLINE TRUNCATED***

**WANINDYATAMI FIRSTIDI PUTRI
NRP 062115 4000 0037**

**Dosen Pembimbing
Jerry Dwi Trijoyo Purnomo, S.Si.,M.Si.,Ph.D**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS SAINS DAN ANALITIKA DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2020**



FINAL PROJECT - KS184822

**MODELING THE PERSENTAGE OF MALNUTRITION
INFANT IN INDONESIA IN 2018 USING
SEMIPARAMETRIC SPLINE TRUNCATED REGRESSION**

**WANINDYATAMI FIRSTIDI PUTRI
SN 062115 4000 0037**

**Supervisor
Jerry Dwi Trijoyo Purnomo, S.Si.,M.Si.,Ph.D**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF SCIENCE AND DATA ANALYTICS
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2020**

LEMBAR PENGESAHAN

**PEMODELAN PERSENTASE BALITA GIZI BURUK
DI INDONESIA TAHUN 2018 MENGGUNAKAN
REGRESI SEMIPARAMETRIK *SPLINE TRUNCATED***

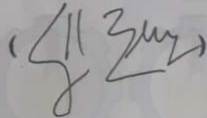
TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Statistika
pada
Program Studi Sarjana Departemen Statistika
Fakultas Sains dan Analitika Data
Institut Teknologi Sepuluh Nopember

Oleh :

Wanindyatami Firstidi Putri
NRP. 062115 4000 0037

Disetujui oleh Pembimbing:
Jerry Dwi Trijoyo P., S.Si., M.Si.,
NIP. 19810223 200812 1 003



Mengetahui,
Kepala Departemen

Dra. Kartika Fithriasari, M.Si
NIP. 19691212 199303 2 002

SURABAYA, JANUARI 2020

PEMODELAN PERSENTASE BALITA GIZI BURUK DI INDONESIA TAHUN 2018 MENGGUNAKAN REGRESI SEMIPARAMETRIK *SPLINE TRUNCATED*

Nama Mahasiswa : Wanindyatami Firstidi Putri
NRP : 062115 4000 0037
Departemen : Statistika-FSAD-ITS
Dosen Pembimbing : Jerry Dwi Trijoyo P, S.Si., M.Si., Ph.D.

Abstrak

Gizi buruk pada balita usia 0-59 bulan merupakan penyakit yang sering terjadi pada balita, karena balita mengalami siklus dan perkembangan yang membutuhkan zat gizi yang lebih banyak dari kelompok umur yang lain. Masalah balita merupakan masalah yang paling serius bagi Negara Indonesia. Penyebab meningkatnya persentase balita gizi buruk di Indonesia dipengaruhi beberapa faktor-faktor, untuk mengetahui faktor yang mempengaruhinya dilakukan penelitian menggunakan Regresi Semiparametrik Spline Truncated. Karena pola hubungan yang ditunjukkan antar persentase balita gizi buruk dengan faktor-faktor yang diduga mempengaruhinya ada yang mengikuti pola parametric dan nonparametrik. Berdasarkan nilai GCV yang paling minimum, model terbaik adalah menggunakan semiparametrik spline truncated linier dengan tiga titik knot. Hasil pengujian signifikansi parameter menunjukkan bahwa seluruh variabel yang digunakan dalam penelitian berpengaruh signifikan terhadap persentase balita gizi buruk di Indonesia. Variabel yang digunakan adalah persentase penduduk miskin dan persentase posyandu aktif. Hasil pengujian asumsi residual menunjukkan semua asumsi terpenuhi dengan nilai koefisien determinasi dari model ini sama dengan 66,94%.

Kata Kunci : Balita Gizi Buruk, GCV, Regresi Semiparametrik, Spline Truncated, Titik Knot

(Halaman ini sengaja dikosongkan)

MODELING PERSENTACE OF MALNUTRITION INFANT IN INDONESIA IN 2018 USING SEMIPARAMETRIC SPLINE TRUNCATED REGRESSION

Name : Wanindyatami Firstidi Putri
Student Number : 062115 4000 0037
Department : Statistics
Supervisor : Jerry Dwi Trijoyo P, S.Si., M.Si., Ph.D.

Abstract

Malnutrition in infants aged 0-59 months is a disease that often occurs in infants, because toddlers experience cycles and development that requires more nutrients than other age groups. The problem of toddlers is the most serious problem for Indonesia. The cause of the increasing percentage of malnourished children under five in Indonesia is influenced by several factors. Because the pattern of relationships shown between the percentage of malnourished children under five with factors thought to influence it is some that follow parametric and nonparametric patterns. Based on the minimum GCV values, the best model is to use a semi-parametric spline truncated linear with three knots. The result of the testing of the significance of the parameters shows that all variables used in the study have a significant effect on the percentage of malnourished children under five in Indonesia. The variables used are the percentage of poor population and the percentage of active posyandu. The results of testing the residual assumptions indicate all assumptions are met with the coefficient of determination of this model equal to 66.94%

Keywords: *Malnutrition Infant, GCV, Semiparametric Regression, Spline Truncated, Knot Points*

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Segala puji syukur penulis panjatkan kepada Allah SWT, Tuhan semesta alam, atas segala rahmat, hidayah, dan karunia-Nya. Atas ridho-Nya penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul **“Pemodelan Persentase Balita Gizi Buruk di Indonesia tahun 2018 Menggunakan Regresi Semiparametrik *Spline Truncated*”**.

Penulis menyadari bahwa penyusunan Tugas Akhir ini tidak dapat terselesaikan tanpa dukungan berbagai pihak baik berupa dukungan moril, dan materil. Oleh karena itu, dengan penuh kerendahan hati, pada kesempatan ini penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Jerry Dwi Trijoyo Purnomo, S.Si., M.Si., Ph.D. selaku dosen pembimbing saya yang telah banyak meluangkan waktunya untuk memberikan bimbingan ajaran, dan ilmu-ilmu baru yang penulis dapatkan dari selama penyusunan Tugas Akhir ini. Terimakasih dan mohon maaf apabila ada kesalahan yang telah penulis lakukan.
2. Ibu Dr. Dra. Kartika Fithriarsari, M.Si. selaku Kepala Departemen Statistika dan Ibu Dr. Santi Wulan P, S.Si, M.Si selaku Sekretaris Departemen 1 yang telah menyediakan fasilitas untuk mendukung kelancaran penyusunan Tugas Akhir.
3. Bapak Prof. Dr. Drs. I Nyoman Budiantara dan Bu Dr. Santi Wulan Purnami, S.Si., M.Si selaku dosen penguji yang telah memberikan saran yang membangun untuk kesempurnaan Tugas Akhir ini.
4. Bapak Dr. Purhadi, M.Sc selaku dosen wali yang telah memberikan pengarahan dan wawasan seputar akademik.
5. Seluruh dosen dan karyawan Departemen Statistika ITS yang telah memberikan ilmu selama penulis menempuh Pendidikan dan membantu kelancaran dan kemudahan dalam pelaksanaan kegiatan perkuliahan.
6. Kedua orang tua yang penulis sayangi, Bapak Budi Santosa dan Ibu Triyanti yang telah mengirimkan do'a yang tidak hentinya dan kasih yang tulus terhadap penulis. Serta adik penulis,

- Satriya Pinandhita Sconditya Putra, yang selalu memberi semangat dan dukungan kepada penulis.
7. Sahabat terbaik penulis selama perkuliahan Azizah, Vienesca Laurencia, dan Nabila Savina. Terima kasih telah memberi semangat, mengisi hari-hari penulis selama masa perkuliahan, dan selalu saling menguatkan.
 8. Sahabat Kost Beta yang selalu satu kost sama penulis. Vathya Rizkiana, Vivien Maulidya terimakasih atas hiburan dan semangat untuk penulis.
 9. Sahabat yang selalu mendengar keluh kesah penulis saat menyelesaikan Tugas Akhir. Indah Novita Kartika Dewi, Wahyu Tri Astuti, Dwi Ningsih, dan Chika Febriyanti. Terimakasih atas semangat, hiburan dan kerecehan selama ini yang telah diberikan.
 10. Dewi Wahyu selaku teman seperjuangan Tugas Akhir S1 Statistika 2015, terimakasih untuk semangat dan diskusi yang diberikan kepada penulis. Semoga penulis dan Dewi bisa menyelesaikan dengan bersama.
 11. Teman-teman Statistika ITS angkatan 2015, Vivacious, yang selalu memberikan dukungan kepada penulis selama ini.
 12. Serta semua pihak yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari bahwa Tugas Akhir ini masih jauh dari sempurna dikarenakan terbatasnya pengalaman dan pengetahuan yang dimiliki penulis. Oleh karena itu, penulis mengharapkan segala bentuk saran serta masukan bahkan kritik yang membangun dari berbagai pihak. Semoga Tugas Akhir ini dapat bermanfaat bagi para pembaca dan semua pihak.

Surabaya, Januari 2020

Penulis

DAFTAR ISI

Halaman

HALAMAN JUDUL	i
COVER PAGE	ii
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR GAMBAR	xvii
DAFTAR TABEL	xix
DAFTAR LAMPIRAN	xxi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan	5
1.4 Manfaat	5
1.5 Batasan Masalah	5
BAB II TINJAUAN PUSTAKA	7
2.1 Statistika Deskriptif	7
2.2 Analisis Regresi	8
2.3 Regresi Nonparametrik	8
2.4 Regresi Semiparametrik	9
2.5 Regresi <i>Spline Truncated</i>	9
2.6 Estimasi Parameter Regresi Nonparametrik <i>Spline Truncated</i>	10
2.7 Pemilihan Titik Knot Optimum	12
2.8 Pengujian Parameter Model Regresi	12
2.8.1 Pengujian Secara Serentak	12
2.8.2 Pengujian Secara Parsial	14
2.9 Kriteria Pemilihan Model Terbaik	15
2.10 Pengujian Asumsi Residual Model Regresi	15

2.10.1	Asumsi Identik	15
2.10.2	Asumsi Independen	16
2.10.3	Asumsi Berdistribusi Normal	17
2.11	Balita Gizi Buru	18
BAB III	METODELOGI PENELITIAN	19
3.1	Sumber Data	19
3.2	Kerangka Konsep	19
3.3	Variabel Penelitian	20
3.4	Struktur Data	23
3.5	Langkah Penelitian	23
3.6	Diagram Alir	24
BAB IV	ANALISIS DAN PEMBAHASAN	27
4.1	Karakteristik Persentase Balita Gizi Buruk dan Faktor-Faktor yang Diduga Mempengaruhinya ...	27
4.2	Pola Hubungan Persentase Balita Gizi Buruk dengan Faktor yang Diduga Mempengaruhinya ...	29
4.3	Pemodelan Persentase Balita Gizi Buruk dengan Regresi Semiparametrik <i>Spline Truncated</i>	31
4.3.1	Pemilihan Titik Knot Optimum Spline Linier	31
4.3.2	Pemilihan Model Terbaik	39
4.4	Penaksir Parameter Model Regresi Semiparametrik <i>Spline Truncated Linier</i>	40
4.5	Pengujian Signifikan	40
4.5.1	Uji Serentak	40
4.5.2	Uji Parsial	41
4.6	Uji Asumsi Residual	42
4.3.1	Asumsi Identik	42
4.3.2	Asumsi Independen	43
4.3.3	Asumsi Distribusi Normal	43
4.7	Koefisien Determinasi	44
4.8	Interpretasi Model <i>Spline</i> Terbaik pada Variabel Persentase Penduduk Miskin (X_2)	45

BAB V KESIMPULAN DAN SARAN	47
5.1 Kesimpulan	47
5.2 Saran	48
DAFTAR PUSTAKA	49
LAMPIRAN	51
BIODATA PENULIS	93

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

	Halaman
Gambar 3.1 Kerangka Konsep	19
Gambar 3.2 Diagram Alir	24
Gambar 4.1 <i>Scatterplot</i> Persentase Balita Gizi Buruk dan Persentase Penduduk Miskin	30
Gambar 4.2 <i>Scatterplot</i> Persentase Balita Gizi Buruk dan Persentase Posyandu Aktif	30
Gambar 4.3 Hasil Uji <i>Kolmogrov Smirnov</i> dan Plot Normalitas Residual	44

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

		Halaman
Tabel 2.1	Analisis ragam (ANOVA) Uji Parameter	13
Tabel 3.1	Variabel Penelitian	20
Tabel 3.2	Struktur Data	23
Tabel 4.1	Karakteristik Data	27
Tabel 4.2	Nilai GCV Regresi Semiparametrik <i>Spline</i> <i>Linier</i> Satu Titik Knot	32
Tabel 4.3	Nilai GCV Regresi Semiparametrik <i>Spline</i> <i>Linier</i> Dua Titik Knot	33
Tabel 4.4	Nilai GCV Regresi Semiparametrik <i>Spline</i> <i>Linier</i> Tiga Titik Knot	34
Tabel 4.5	Nilai GCV Regresi Semiparametrik <i>Spline</i> <i>Kuadratik</i> Satu Titik Knot	35
Tabel 4.6	Nilai GCV Regresi Semiparametrik <i>Spline</i> <i>Kuadratik</i> Dua Titik Knot	35
Tabel 4.7	Nilai GCV Regresi Semiparametrik <i>Spline</i> <i>Kuadratik</i> Tiga Titik Knot	36
Tabel 4.8	Nilai GCV Regresi Semiparametrik <i>Spline</i> <i>Kubik</i> Satu Titik Knot	37
Tabel 4.9	Nilai GCV Regresi Semiparametrik <i>Spline</i> <i>Kubik</i> Dua Titik Knot	38
Tabel 4.10	Nilai GCV Regresi Semiparametrik <i>Spline</i> <i>Kubik</i> Tiga Titik Knot	39
Tabel 4.11	Perbandingan Nilai GCV Minimum tiap Titik Knot	40
Tabel 4.12	ANOVA Model Regresi Semiparametrik <i>Spline</i>	41
Tabel 4.13	Hasil Pengujian Signifikan Parameter Secara Individu	41
Tabel 4.14	Hasil Pengujian Statistik Uji <i>Glejser</i>	42
Tabel 4.15	Uji Asumsi Residual Independen	43

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

	Halaman
Lampiran 1.	Data Persentase Kasus HIV di Indonesia Tahun 2017 dengan Faktor-faktor yang 51
Lampiran 2.	<i>Scatter plot</i> Variabel Y terhadap variabel X_1 , X_2 , X_3 , X_4 , dan X_5 52
Lampiran 3.	<i>Syntax</i> Pemilihan Titik Knot Optimal dengan Linier Satu Titik Knot Menggunakan R 53
Lampiran 4.	<i>Syntax</i> Pemilihan Titik Knot Optimal dengan Linier Dua Titik Knot Menggunakan R 55
Lampiran 5.	<i>Syntax</i> Pemilihan Titik Knot Optimal dengan Linier Tiga Titik Knot Menggunakan R 57
Lampiran 6.	<i>Syntax</i> Pemilihan Titik Knot Optimal dengan Kuadratik Satu Titik Knot Menggunakan R 60
Lampiran 7.	<i>Syntax</i> Pemilihan Titik Knot Optimal dengan Kuadratik Dua Titik Knot Menggunakan R 62
Lampiran 8.	<i>Syntax</i> Pemilihan Titik Knot Optimal dengan Kuadratik Tiga Titik Knot Menggunakan R 64
Lampiran 9.	<i>Syntax</i> Pemilihan Titik Knot Optimal dengan Kubik Satu Titik Knot Menggunakan R 67
Lampiran 10.	<i>Syntax</i> Pemilihan Titik Knot Optimal dengan Kubik Dua Titik Knot Menggunakan R 69
Lampiran 11.	<i>Syntax</i> Pemilihan Titik Knot Optimal dengan Kubik Tiga Titik Knot Menggunakan R 72

Lampiran 12.	Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Linier Satu Titik Knot	75
Lampiran 13.	Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Linier Dua Titik Knot	76
Lampiran 14.	Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Linier Tiga Titik Knot	77
Lampiran 15.	Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Kuadratik Satu Titik Knot	78
Lampiran 16.	Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Kuadratik Dua Titik Knot	79
Lampiran 17.	Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Kuadratik Tiga Titik Knot	80
Lampiran 18.	Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Kubik Satu Titik Knot	81
Lampiran 19.	Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Kubik Dua Titik Knot.....	82
Lampiran 20.	Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Kubik Tiga Titik Knot	83
Lampiran 21.	<i>Syntax</i> Estimasi dan Pengujian Parameter Menggunakan R	84
Lampiran 22.	<i>Output</i> Estimasi dan Pengujian Parameter Menggunakan R	87
Lampiran 23.	<i>Output</i> Residual Menggunakan R	88
Lampiran 24.	<i>Syntax</i> Uji Glejser Menggunakan R	89
Lampiran 25.	<i>Output</i> Uji Glejser Menggunakan R	91
Lampiran 30.	Surat Keterangan Pengambilan Data	92

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan salah satu negara berkembang yang berkomitmen untuk berkontribusi terhadap tujuan dan target *Sustainable Development Goals*. *Sustainable Development Goals* (SDGs) sebagai kesepakatan pembangunan global, yang dirancang oleh 191 negara Perserikatan Bangsa-Bangsa (PBB) yang hadir. *Sustainable Development Goals* 2015-2030 secara resmi menggantikan *Millennium Development Goals* (MDGs). SDGs memiliki 17 tujuan dan 169 target yang harus dicapai, salah satu tujuan SDGs adalah untuk mengatur hidup sehat dan meningkatkan kesejahteraan untuk semua usia. Salah satu target yang ingin dicapai pada tujuan tersebut adalah tentang gizi masyarakat yang diharapkan dapat mengakhiri segala bentuk malnutrisi, termasuk mencapai target internasional 2025 untuk penurunan *stunting* dan *wasting* pada balita. Hasil Sensus menunjukkan adanya penurunan persentase gizi buruk di Indonesia pada tahun 2013 yaitu sebanyak 5,7% dan pada tahun 2018 menurun menjadi 3,9% (Kementerian kesehatan Republik Indonesia, 2019), akan tetapi target SDGs belum tercapai. Pengukuran status gizi didasarkan atas standar *World Health Organization* (WHO, 2005) yang telah ditetapkan pada keputusan Menteri Kesehatan Nomor 1995/Menkes/SK/XII/-2010 tentang Standar Antropometri Penilaian Status Gizi Anak. Menurut standar tersebut, status gizi balita dapat diukur berdasarkan berat badan menurut umur (BB/U) (Kementerian kesehatan Republik Indonesia, 2019).

Balita merupakan kelompok masyarakat yang rentan gizi karena mengalami siklus dan perkembangan yang membutuhkan zat gizi yang lebih banyak dari kelompok umur yang lain sehingga balita paling mudah menderita status gizi (Kementerian kesehatan Republik Indonesia, 2019). Status gizi terdiri dari dua status yaitu gizi buruk, dan gizi kurang. Status gizi balita hal yang sangat penting yang harus diketahui oleh setiap orang tua, maka status gizi

tersebut perlu dideteksi sejak dini. Masalah balita gizi buruk merupakan masalah yang sangat serius, apabila tidak ditangani secara cepat dapat menimbulkan pengaruh yang sangat menghambat pertumbuhan fisik, mental maupun kemampuan berpikir yang bisa membuat produktifitas kerja menurun. Balita penderita gizi buruk dapat mengalami penurunan (IQ) hingga sepuluh persen, selain itu juga, penyakit yang dapat diderita sama balita gizi buruk adalah diabetes (kencing manis) dan penyakit jantung koroner. Dampak yang paling buruk terhadap balita gizi buruk ini adalah kematian pada usia yang sangat dini (Rahim, 2014).

Gizi buruk pada balita usia 0-59 bulan, hasil Riset Kesehatan Dasar tahun 2018 menyatakan bahwa persentase gizi buruk di Indonesia adalah 3,9% hal ini tidak berbeda jauh dengan hasil Pemantauan Status Gizi (PSG) yang diselenggarakan oleh Kementerian Kesehatan tahun 2017 yaitu persentase balita gizi buruk sebesar 3,8%. Provinsi dengan persentase tertinggi gizi buruk pada balita usia 0-59 bulan tahun 2018 adalah Nusa Tenggara, sedangkan provinsi dengan persentase rendah adalah Provinsi Kepulauan Riau (Kementerian kesehatan Republik Indonesia, 2019). Dari hasil riset tersebut gizi buruk masih belum mencapai target SDGs, maka perlu dilakukan upaya supaya target SDGs tercapai, sehingga gizi buruk pada balita berkurang maka balita bisa hidup sehat dan pertumbuhan balita bisa tumbuh kembang tidak terhambat pertumbuhannya fisiknya. Salah satu penanganan yang cepat dan tepat sesuai tata laksana kasus anak gizi buruk adalah melakukan penimbangan balita secara rutin sehingga balita dapat dipantau secara intensif. Hal ini dimaksud apabila berat badan anak tidak naik atau jika ditemukan penyakit, dapat segera dilakukan upaya pemulihan dan pencegahan, agar tidak terjadi gizi buruk (Kementerian kesehatan Republik Indonesia, 2019). Banyak faktor yang mempengaruhi terjadinya gizi buruk pada balita usia 0-59 tahun, diantaranya adalah status sosial, ekonomi, ketidaktahuan ibu tentang pemberian gizi yang

baik untuk anak, dan Berat Badan Lahir Rendah (BBLR) (Kusriadi, 2010).

Penelitian mengenai gizi buruk di Indonesia masih sedikit, penelitian sebelumnya yang telah dilakukan oleh Novitasari (2012) menyatakan bahwa faktor yang mempengaruhi gizi buruk pada anak balita adalah status social ekonomi, pendidikan ibu, penyakit komorbiditas, ASI, berat badan lahir rendah, dan kelengkapan imunisasi. Faktor resiko gizi buruk yang paling dominan adalah anak-anak dibawah lima tahun penyakit komorbiditas. Kemudian Saputra dan Nurriszka (2012) membahas masalah gizi buruk dan gizi kurang pada tiga komunitas di Sumatra Barat, yaitu komunitas perkotaan, komunitas perikanan, dan komunitas pertanian melalui studi terhadap 572 keluarga yang dipilih dengan menggunakan teknik *purposive random sampling* dan dalam penelitian tersebut disebutkan bahwa kemiskinan dan tingkat pendidikan orang tua merupakan faktor utama penyebab balita gizi buruk dan gizi kurang. Selanjutnya penelitian dari Mazarina Devi (2010), asupan gizi yang baik sering tidak bias terpenuhi oleh seorang anak, diantaranya karena faktor ekonomi keluarga, pendidikan, dan jumlah keluarga. Faktor yang menjadi variabel independen terdiri dari umur anak, jenis kelamin anak, usia orang tua, tingkat pendidikan orang tua, pekerjaan orang tua, jumlah anggota keluarga, dan lama menyusui. Pada penelitian tersebut faktor yang paling dominan atau berhubungan dengan status gizi adalah jenis pekerjaan ayah dan jenis pekerjaan ibu. Ketiga penelitian tersebut merupakan referensi dengan kasus yang serupa dan metode yang berbeda, sehingga penelitian tersebut belum memperhatikan mengenai pola data yang tidak terdeteksi antara variabel prediktor dan variabel respon, maka penelitian tersebut perlu dilakukan pemodelan menggunakan regresi nonparametrik (Eubank, 1999).

Penelitian ini bertujuan untuk memodelkan faktor-faktor yang mempengaruhi persentase balita gizi buruk di Indonesia dapat dilakukan menggunakan beberapa metode, salah satunya adalah analisis regresi. Analisis regresi terdapat tiga pendekatan yaitu

regresi parametrik, regresi nonparametrik, dan regresi semi-parametrik (Budiantara, 2007). Pada penelitian ini menggunakan data persentase balita gizi buruk di Indonesia tahun 2018 sebagai variabel dependen dan sebagai variabel independen adalah faktor-faktor yang diduga mempengaruhi balita gizi buruk diantaranya adalah bayi mendapat ASI eksklusif, penduduk miskin, kabupaten/kota yang melaksanakan PHBS, posyandu aktif, dan kunjungan neonatal pertama. Metode regresi yang digunakan adalah regresi semiparametrik *Spline Truncated*, karena pada penelitian ini terdapat beberapa variabel dependen yang memiliki hubungan linier terhadap variabel independen, sedangkan beberapa variabel independen lain memiliki hubungan non linier dengan variabel dependen. Salah satu pengembangan regresi semiparametrik adalah *spline*. Metode regresi semiparametrik *spline* dipilih karena memiliki kelebihan yaitu metode memiliki fleksibilitas tinggi dan dapat memodelkan data yang memiliki pola berubah-ubah pada interval tertentu. Kelebihan ini terjadi karena dalam *spline* terdapat titik-titik knot, yaitu titik perpaduan bersama yang menunjukkan terjadinya perubahan pola perilaku data (Budiantara, 2007). Penelitian ini diharapkan menjadi bahan pertimbangan atau rekomendasi bagi pemerintah Indonesia untuk menekan persentase balita gizi buruk dan meningkatkan penanganan balita gizi buruk di Indonesia tiap tahunnya dan penelitian ini dapat memberikan pemodelan yang terbaik untuk data persentase balita gizi buruk di Indonesia tahun 2018, sehingga dapat diketahui pola hubungannya terhadap faktor-faktor yang diduga mempengaruhinya.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan mengenai persentase balita gizi buruk di Indonesia, maka rumusan masalah yang akan diselesaikan dalam penelitian ini adalah sebagai berikut.

1. Bagaimana karakteristik dan faktor-faktor yang mempengaruhi persentase balita gizi buruk di Indonesia tahun 2018?

2. Bagaimana pemodelan persentase balita gizi buruk di Indonesia tahun 2018 menggunakan metode Regresi Semi-parametrik *Spline Truncated*?

1.3 Tujuan

Berdasarkan rumusan masalah yang telah disebutkan, tujuan yang ingin dicapai pada penelitian ini adalah sebagai berikut.

1. Mengetahui karakteristik dan faktor-faktor yang mempengaruhi persentase Balita Gizi Buruk di Indonesia pada tahun 2018.
2. Memodelkan data persentase Balita Gizi Buruk di Indonesia tahun 2018 menggunakan Regresi Semiparametrik *Spline Truncated*.

1.4 Manfaat

Manfaat yang diharapkan dari hasil penelitian ini adalah dapat dijadikan bahan acuan bagi pihak pemerintah Indonesia mengenai persentase balita gizi buruk di Indonesia sehingga dapat dijadikan masukan dalam pengambilan kebijakan. Bagi pembaca, penelitian ini diharapkan mampu menerapkan ilmu yang sejalan dengan bidang Statistika yaitu metode regresi semiparametrik *spline* dalam bentuk nyata khususnya dalam bidang sosial pemerintah, serta dapat bermanfaat sebagai referensi bagi pembaca dan peneliti selanjutnya.

1.5 Batasan Masalah

Batasan masalah pada penelitian ini adalah penelitian menggunakan data persentase balita gizi buruk di Indonesia pada tahun 2018 dengan unit observasi 34 Provinsi di Indonesia. Fungsi yang digunakan adalah *Spline* dengan titik knot yang digunakan dalam pemodelan adalah satu, dua, dan tiga titik knot. Pemilihan titik knot optimal pada penelitian ini menggunakan *Generalized Cross Validation (GCV)*..

(Halaman ini sengaja dikosongkan)

BAB II

TINJAUAN PUSTAKA

2.1 Statistika Deskriptif

Statistika deskriptif merupakan metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna. Metode ini bertujuan untuk menguraikan tentang sifat-sifat dari suatu keadaan dan membuat deskripsi atau gambaran yang sistematis dan akurat mengenai fakta-fakta dan sifat-sifat dari fenomena yang diselidiki. Ukuran pemusatan data dan ukuran penyebaran data merupakan suatu alat yang dapat digunakan untuk mendefinisikan ukuran-ukuran numerik yang menjelaskan karakteristik dari data (Walpole, 1993). Statistika deskriptif dapat menyajikan suatu data dalam bentuk tabel, diagram batang, histogram, *scatter plot* sehingga informasi yang disampaikan akan lebih mudah dipahami oleh pembaca.

Ukuran pemusatan data, bisa menggunakan rata-rata (*mean*) dan median. Rata-rata adalah hasil pembagian antara jumlah nilai setiap pengamatan dengan banyaknya data pengamatan (Walpole, 1993).

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (2.1)$$

dengan:

\bar{x} : *mean*

x_i : pengamatan ke- i , $i=1,2,\dots,n$

n : banyak pengamatan.

Untuk ukuran penyebaran data, biasa digunakan varians (s^2). Varians adalah kuadrat simpangan dari semua nilai data terhadap rata-rata dan dinyatakan dengan simbol σ^2 untuk populasi dan s^2 untuk sampel (Walpole, 1993).

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (2.2)$$

Selain itu, dalam statistika deskriptif yang sering dianalisis adalah nilai maksimum dan nilai minimum. Nilai maksimum merupakan nilai tertinggi/terbesar yang terdapat dalam suatu kelompok data. Sedangkan nilai minimum adalah nilai terendah yang terdapat dalam sekumpulan data.

2.2 Analisis Regresi

Analisis regresi merupakan salah satu metode statistik yang menjelaskan tentang pola hubungan (model) antara dua variabel atau lebih. Variabel-variabel tersebut dikelompokkan menjadi variabel respon yang dinotasikan dengan huruf y dan variabel prediktor dinotasikan dengan huruf x (Draper dan Smith, 1998). Pola hubungan antara variabel respon dengan variabel prediktor dapat diketahui melalui *scatterplot*, bentuk pola dapat membentuk pola linier, kuadratik, kubik, atau acak. Dalam analisis regresi terdapat tiga model yaitu model regresi parametrik, model regresi semiparametrik, dan model regresi nonparametrik. Pada model regresi parametrik bentuk kurva regresi diasumsikan diketahui dan diperlukan pengetahuan masa lalu tentang karakteristik data yang akan diselidiki. Sedangkan regresi nonparametrik bentuk kurva regresi diasumsikan tidak diketahui. Kurva regresi nonparametrik hanya diasumsikan *smooth* (mulus) dalam arti termuat di dalam suatu fungsi tertentu. Bentuk estimasi dari data tanpa diketahui oleh faktor subyektifitas peneliti. Regresi semiparametrik digunakan jika di dalam suatu model regresi terdapat komponen parametrik dan nonparametrik (Budiantara, 2001).

2.3 Regresi Nonparametrik

Regresi nonparametrik merupakan model regresi yang digunakan untuk mengetahui adanya pola hubungan antara variabel respon dengan variabel prediktor yang tidak diketahui

bentuk kurva regresinya. Dengan kata lain regresi nonparametrik memiliki fleksibilitas yang tinggi dalam memodelkan pola data (Eubank, 1999). Model regresi nonparametrik dapat ditulis pada persamaan (2.3) sebagai berikut.

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.3)$$

dimana y_i adalah variabel respon ke- i , $f(x_i)$ adalah nilai dari fungsi regresi yang tidak diketahui pada titik $x_1, x_2, \dots, x_n, x_i$ adalah variabel prediktor, dan ε_i merupakan error dengan asumsi normal, independen, dan identik dengan mean nol dan variansi σ^2 .

2.4 Regresi Semiparametrik

Regresi semiparametrik merupakan gabungan antara regresi parametrik dan regresi nonparametrik (Carrol, Wand, dan Ruppert, 2003). Analisis regresi semiparametrik ini digunakan ketika bentuk pola hubungan antara variabel prediktor dan variabel respon sebagian diketahui bentuknya dan sebagian lagi tidak diketahui bentuknya. Berikut ini merupakan model regresi semiparametrik:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + f(x_i) + \varepsilon_i \quad (2.4)$$

Dimana $i = 1, 2, \dots, n$ dan x_1, x_2, \dots, x_p merupakan variabel prediktor untuk komponen parametrik dan x_i sebagai variabel prediktor untuk komponen nonparametrik.

2.5 Regresi *Spline Truncated*

Salah satu model regresi semiparametrik yang digunakan adalah *Spline*. *Spline* merupakan bentuk kurva yang terpotong-potong sehingga mampu mengatasi perubahan pola data pada sub interval tertentu. Pada metode regresi semiparametrik *spline truncated* digunakan bantuan titik-titik knot. Titik knot merupakan titik dimana terjadi pola perubahan perilaku dari suatu fungsi pada selang yang berbeda (Hardle, 1990). Regresi *spline truncated*

merupakan potongan polynomial yang fleksibel. Fungsi *spline truncated* diperoleh dari hasil penjumlahan antara fungsi polinomial dengan fungsi *truncated*. Misal fungsi *spline truncated* berorde p dengan titik knot K_1, K_2, \dots, K_r . Kurva regresi yang terbentuk adalah $f(x_i)$, lebih rinci dapat dituliskan menjadi persamaan sebagai berikut :

$$f(x_i) = \sum_{j=0}^p \beta_j x_i^j + \sum_{k=1}^r \beta_{p+k} (x_i - K_k)_+^p \quad (2.5)$$

dengan fungsi *truncated* $(x_i - K_k)_+^p$ akan menghasilkan persamaan (2.6) berikut,

$$(x_i - K_k)_+^p = \begin{cases} (x_i - K_k)^p, & x_i \geq K_k \\ 0 & , x_i < K_k \end{cases} \quad (2.6)$$

Titik $x = K_k$ adalah titik knot yang menggambarkan pola perubahan fungsi pada sub interval yang berbeda dan nilai p adalah derajat polinomial (Eubank, 1999).

2.6 Estimasi Parameter Regresi Nonparametrik *Spline Truncated*

Estimasi parameter pada regresi nonparametrik *spline truncated* menggunakan metode *Ordinary Least Square* (OLS). Metode ini mengestimasi parameter dengan cara meminimumkan jumlah kuadrat residual. Berikut adalah bentuk matriks dari model regresi semiparametrik *spline* linear dengan r knot dan univariabel prediktor.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.7)$$

dimana

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{X} = \begin{pmatrix} 1 & x_1 & (x_1 - k_1)_+^1 & \cdots & (x_1 - k_r)_+^1 \\ 1 & x_2 & (x_2 - k_1)_+^1 & \cdots & (x_2 - k_r)_+^1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & (x_n - k_1)_+^1 & \cdots & (x_n - k_r)_+^1 \end{pmatrix},$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_r \end{pmatrix}, \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Dari persamaan (2.7), persamaan residual dapat ditulis menjadi bentuk persamaan berikut,

$$\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta} \quad (2.8)$$

Jumlah kuadrat residual yang berupa matriks dapat ditulis sebagai berikut,

$$\begin{aligned} \sum_{i=1}^n \varepsilon_i^2 &= \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} \\ &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &= \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \\ &= \mathbf{y}'\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \end{aligned} \quad (2.9)$$

Untuk meminimumkan $\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}$ maka turunan pertama terhadap $\boldsymbol{\beta}$ harus sama dengan nol.

$$\frac{\partial(\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon})}{\partial\boldsymbol{\beta}} = 0 \quad (2.10)$$

Kemudian didapatkan persamaan sebagai berikut:

$$\begin{aligned} -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} &= 0 \\ \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} &= \mathbf{X}'\mathbf{y} \\ (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ \hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \end{aligned} \quad (2.11)$$

2.7 Pemilihan Titik Knot Optimum

Model regresi *spline* terbaik merupakan model yang memiliki titik knot optimal. Titik knot merupakan titik perpaduan bersama dimana terjadi pola perubahan perilaku dari suatu fungsi pada selang yang berbeda (Hardle, 1990). Salah satu metode yang biasa digunakan untuk memilih titik knot optimal adalah metode *Generalized Cross Validation* (GCV). Metode GCV memiliki beberapa kelebihan diantaranya mempunyai sifat optimal asimtotik, tidak memuat varians populasi (σ^2) yang tidak diketahui, *invariance* terhadap transformasi. Model regresi *spline* terbaik diperoleh dari titik knot optimal dengan melihat nilai GCV terkecil (Wahba, 1990). Metode GCV dapat dituliskan sebagai berikut.

$$GCV(K_1, K_2, \dots, K_r) = \frac{MSE(K_1, K_2, \dots, K_r)}{[n^{-1} \text{trace}(\mathbf{I} - \mathbf{A})]^2} \quad (2.12)$$

dimana \mathbf{I} merupakan matriks identitas, n adalah jumlah pengamatan, $K = (K_1, K_2, \dots, K_r)$ merupakan titik-titik knot,

$$MSE(K_1, K_2, \dots, K_r) = n^{-1} \sum_{i=1}^n (y_i - \hat{f}(x_i))^2 \quad (2.13)$$

serta $\mathbf{A} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ (Eubank, 1999).

2.8 Pengujian Parameter Model Regresi

Uji parameter model regresi dilakukan untuk mengetahui apakah variabel prediktor memiliki pengaruh signifikan terhadap variabel respon atau tidak. Pada regresi semiparametrik *spline truncated*, pengujian parameter model regresi dilakukan setelah mendapatkan model regresi dengan titik knot optimal berdasarkan nilai GCV yang paling minimum. Pengujian parameter dapat dilakukan secara serentak dan individu (parsial) dengan penjelasan pada 2.8.1 dan 2.8.2 sebagai berikut.

2.8.1 Pengujian Secara Serentak

Pengujian secara serentak adalah uji untuk mengetahui apakah semua variable predictor yang terdapat dalam model

memberikan pengaruh secara bersama-sama atau tidak. Hipotesis untuk pengujian parameter secara serentak adalah sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{p+r} = 0$$

$$H_1 : \text{minimal ada satu } \beta_j \neq 0; j = 1, 2, \dots, p+r$$

Dimana p adalah jumlah variabel prediktor dan r adalah jumlah titik knot. Statistik uji dalam pengujian serentak menggunakan uji F seperti pada persamaan 2.14 berikut ini.

$$F_{hitung} = \frac{MS_{regresi}}{MS_{residual}} \quad (2.14)$$

Uji parameter model secara serentak dapat disajikan menggunakan *Analysis of Varians* (ANOVA) yang disajikan dalam tabel 2.1 sebagai berikut.

Tabel 2.1 Analisis ragam (ANOVA) Uji Parameter

Sumber Variasi	Df	Sum of Square (SS)	Mean Square (MS)	F_{hitung}
Regresi	$p+r$	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$\frac{SS_{regresi}}{df_{regresi}}$	$\frac{MS_{regresi}}{MS_{error}}$
Error	$n-(p+r)-1$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{SS_{error}}{df_{error}}$	
Total	$n-1$	$\sum_{i=1}^n (y_i - \bar{y})^2$		

Nilai $p+r$ merupakan banyak parameter dalam model regresi semiparametrik *spline* kecuali β_0 . Tolak H_0 jika $F_{hitung} > F_{\alpha; (p+r, n-(p+r)-1)}$ atau $p\text{-value} < \alpha$ sehingga menghasilkan kesimpulan bahwa minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon. Untuk itu harus

dilanjutkan pengujian secara parsial yang berfungsi untuk mengetahui variabel-variabel prediktor yang berpengaruh secara signifikan (Draper dan Smith, 1998).

2.8.2 Pengujian Secara Parsial

Pengujian secara individu atau parsial dilakukan apabila pada uji secara serentak didapatkan kesimpulan bahwa minimal terdapat satu parameter yang signifikan. Berikut merupakan hipotesis untuk pengujian secara parsial:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, \dots, p + r$$

Uji parsial menggunakan uji t . Statistik uji yang digunakan adalah sebagai berikut:

$$t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (2.15)$$

$SE(\hat{\beta}_j)$ adalah *standart error* $\hat{\beta}_j$ dimana $SE(\hat{\beta}_j) = \sqrt{Var(\hat{\beta}_j)}$ dengan $Var(\hat{\beta}_j)$ merupakan elemen diagonal utama ke- j dari matriks $Var(\hat{\beta})$, $j = 1, 2, \dots, p + r$ dari matriks yang dapat diurai seperti berikut,

$$\begin{aligned} Var(\hat{\beta}) &= \text{var} \left[(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y} \right] \\ &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \text{var}(\mathbf{Y}) \left[(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \right]' \\ &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' (\sigma^2 \mathbf{I}) \mathbf{X} (\mathbf{X}'\mathbf{X})^{-1} \\ &= \sigma^2 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \mathbf{X} (\mathbf{X}'\mathbf{X})^{-1} \\ &= \sigma^2 (\mathbf{X}'\mathbf{X})^{-1} \end{aligned} \quad (2.16)$$

Tolak H_0 jika $|t_{hitung}| > t_{\frac{\alpha}{2};(n-(p+r)-1)}$ atau $p\text{-value} < \alpha$,

yang dapat disimpulkan bahwa parameter berpengaruh secara signifikan terhadap model (Draper dan Smith, 1998). Pengujian Parameter Model Regresi

2.9 Kriteria Pemilihan Model Terbaik

Salah satu kriteria yang digunakan dalam menentukan model regresi terbaik adalah koefisien determinasi (R^2). Koefisien determinasi merupakan kuantitas yang dapat menjelaskan sumbangan variabel prediktor terhadap variabel respon. Semakin tinggi nilai R^2 yang dihasilkan suatu model, maka semakin baik pula variabel-variabel prediktor dalam model tersebut dalam menjelaskan variabilitas variabel respon (Draper dan Smith, 1998). Berikut ini adalah rumus untuk mendapatkan nilai R^2 .

$$R^2 = \frac{SS_{Regresi}}{SS_{total}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.17)$$

Pemilihan model juga akan menunjukkan banyaknya parameter yang digunakan dalam model tersebut. Seperti yang dijelaskan dalam prinsip parsimoni, suatu model regresi yang baik adalah model regresi dengan banyak parameter yang sesedikit mungkin tetapi mempunyai R^2 yang cukup tinggi.

2.10 Pengujian Asumsi Residual Model Regresi

Pengujian asumsi residual (*goodness of fit*) model regresi dilakukan guna mengetahui apakah residual yang dihasilkan telah memenuhi asumsi identik, independen, dan berdistribusi normal.

2.10.1 Asumsi Identik

Asumsi identik (homoskedastisitas) yang berarti bahwa varians pada residual adalah sama atau identik. Kebalikanannya

adalah kasus heteroskedastisitas, yaitu jika kondisi varians *residual* tidak identik (Gujarati D. , 2003).

$$\text{var}(y_i) = \text{var}(\varepsilon_i) = \sigma^2 ; i = 1, 2, \dots, n \quad (2.18)$$

Uji identik dapat menggunakan uji Glejser. Perumusan hipotesisnya adalah sebagai berikut:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{Minimal ada satu } \sigma_i^2 \neq \sigma^2 ; i = 1, 2, \dots, n$$

Statistik uji yang digunakan adalah,

$$F_{hitung} = \frac{\sum_{i=1}^n \left(|\varepsilon_i| - |\bar{\varepsilon}| \right)^2}{\sum_{i=1}^n \left(|\varepsilon_i| - |\hat{\varepsilon}_i| \right)^2} \quad (2.19)$$

$n - v - 1$

dimana nilai v menunjukkan banyaknya parameter model Glejser dan untuk model regresi semiparametrik *Spline Truncated* seperti persamaan (2.4) dan nilai $v = p + r$.

Tolak H_0 jika $F_{hitung} > F_{\alpha; (v, n-v-1)}$ atau $p\text{-value} < \alpha$ yang berarti bahwa tidak terindikasi terdapat kasus homoskedastisitas dan sebaliknya jika $F_{hitung} < F_{\alpha; (v, n-v-1)}$ atau $p\text{-value} > \alpha$ maka gagal tolak H_0 yang berarti bahwa terindikasi terdapat kasus heteroskedastisitas.

2.10.2 Asumsi Independen

Asumsi independen dilakukan untuk memastikan bahwa tidak terdapat korelasi antar residual atau autokorelasi. Uji yang digunakan untuk mendeteksi kasus autokorelasi adalah uji *Durbin-Watson* (Draper & Smith, 1998). Perumusan hipotesis yang digunakan adalah sebagai berikut.

$$H_0 : \rho = 0 \text{ (tidak terjadi autokorelasi)}$$

$$H_1 : \rho \neq 0 \text{ (terjadi autokorelasi)}$$

Statistik uji:

$$d = \frac{\sum_{i=1}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (2.20)$$

Langkah selanjutnya adalah membandingkan hasil statistik uji dengan tabel Dubin Watson (DW). Tabel DW terdiri atas batas bawah (d_L) dan batas atas (d_U). Berikut beberapa keputusan setelah membandingkan dengan tabel DW.

$d < d_L$: tolak H_0

$d > 4 - d_L$: tolak H_0

$d_U < d < 4 - d_L$: gagal tolak H_0

$d_L \leq d \leq d_U$ dan $4 - d_U \leq d \leq 4 - d_L$: Pengujian tidak meyakinkan

2.10.3 Asumsi Berdistribusi Normal

Pengujian asumsi berdistribusi normal dilakukan untuk mengetahui apakah residual suatu data telah mengikuti distribusi normal atau tidak dengan *mean* nol dan varians σ^2 . Uji yang digunakan adalah uji *Kolmogorov-Smirnov* (Daniel, 1989).

Hipotesis :

$$H_0 : F_n(\varepsilon) = F_0(\varepsilon)$$

$$H_1 : F_n(\varepsilon) \neq F_0(\varepsilon)$$

atau

H_0 : residual berdistribusi normal

H_1 : residual tidak berdistribusi normal

Statistik uji :

$$D = \sup_{\varepsilon} |F_n(\varepsilon) - F_0(\varepsilon)| \quad (2.21)$$

Tolak H_0 apabila $D > D_\alpha$ atau p-value $< \alpha$.

D_α adalah nilai kritis untuk uji *Kolmogorov-Smirnov* satu sampel, diperoleh dari tabel *Kolmogorov Smirnov* satu sampel, $F_n(\varepsilon)$

merupakan nilai peluang kumulatif (fungsi distribusi kumulatif) berdasarkan data sampel, $F_0(\mathcal{E})$ adalah nilai peluang kumulatif (fungsi distribusi kumulatif) dibawah H_0 . Tolak H_0 jika $t > t_{\frac{\alpha}{2};(n-(p+r)-1)}$ dan $t < -t_{\frac{\alpha}{2};(n-(p+r)-1)}$ atau $p\text{-value} < \alpha$. Sehingga menghasikan kesimpulan bahwa variabel prediktor ke- n berpengaruh signifikan terhadap variabel respon (Gujarati D. , 2004).

2.11 Balita Gizi Buruk

Gizi buruk adalah suatu kondisi dimana seseorang dinyatakan kekurangan nutrisi yang ditandai dengan berat badan anak jauh dibawah rata-rata (Dhianasih, 2014). Pengukuran status gizi didasrkan atas Standar *World Health Organization* (WHO, 2005) yang telah ditetapkan pada Keputusan Menteri Kesehatan Nomor 1995/Menkes/SK/XII/2010 tentang Standar Antropometri Penilaian Status Gizi Anak. Menurut standar tersebut, status gizi dapat diukur berdasarkan indeks yaitu berat badan menurut umur (BB/U). Gizi buruk yang dialami oleh balita umur 0-59 bulan dapat menimbulkan pengaruh yang sangat menghambat pertumbuhan fisik, mental maupun kemampuan berpikir yang pada akhirnya dapat menurunkan produktivitas kerja. Balita penderita gizi buruk dapat mengalami penurunan (IQ) hingga sepuluh persen, selain itu juga, penyakit yang dapat diderita sama balita gizi buruk adalah diabetes (kencing manis) dan penyakit jantung koroner. Dampak yang paling buruk terhadap balita gizi buruk ini adalah kematian pada usia yang sangat dini (Rahim, 2014).

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Sumber data yang digunakan pada penelitian ini merupakan data sekunder yang diperoleh dari Profil Kesehatan Indonesia tahun 2018 (Kementerian kesehatan Republik Indonesia, 2019). Data tersebut mengenai balita gizi buruk menurut provinsi di Indonesia tahun 2018. Unit observasi yang digunakan adalah 34 provinsi di Indonesia.

3.2 Kerangka Konsep

Kerangka konsep merupakan suatu bentuk kerangka berpikir yang dapat digunakan sebagai pendekatan dalam memecahkan masalah. Pada penelitian ini kerangka konsep penelitian untuk menentukan faktor-faktor yang diduga berpengaruh terhadap persentase balita gizi buruk yang digunakan untuk penelitian. Berikut merupakan kerangka konsep pada penelitian ini.



Gambar 3.1 Kerangka Konsep

Berdasarkan tiga faktor pada Gambar 3.1 yang mendasari variabel penelitian pada penelitian persentase balita gizi buruk. Faktor tersebut diperoleh variabel penelitian yang diduga berpengaruh terhadap balita gizi buruk. Pada faktor lingkungan diperoleh variabel penduduk miskin dan melaksanakan PHBS, faktor 2 yaitu perilaku diperoleh variabel bayi mendapat ASI eksklusif, dan yang terakhir faktor ke 3 (pelayanan kesehatan) diperoleh variabel posyandu aktif,

dan variabel kunjungan neonatal pertama. Sehingga pada penelitian ini didapatkan 5 variabel prediktor dari 3 faktor tersebut.

3.3 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri dari variabel respon yang digunakan dalam penelitian ini adalah persentase balita gizi buruk di Indonesia tahun 2018 dan beberapa variabel prediktor yang merupakan faktor-faktor yang diduga berpengaruh terhadap persentase balita gizi buruk yang diperoleh dari penelitian-penelitian terdahulu. Variabel penelitian yang digunakan dijelaskan pada Tabel 3.1 berikut.

Tabel 3.1 Kebutuhan Data Penelitian

No.	Variabel	Keterangan
1	Y	Persentase balita gizi buruk di Indonesia menurut Provinsi Tahun 2018
2	X ₁	Persentase bayi mendapat ASI eksklusif di Indonesia menurut Provinsi Tahun 2018
3	X ₂	Persentase penduduk miskin di Indonesia menurut Provinsi Tahun 2018
4	X ₃	Persentase kabupaten/kota yang melaksanakan PHBS di Indonesia menurut Provinsi Tahun 2018
5	X ₄	Persentase posyandu aktif di Indonesia menurut Provinsi Tahun 2018
6	X ₅	Persentase cakupan kunjungan neonatal pertama di Indonesia menurut Provinsi Tahun 2018

Berikut defenisi operasional dari tiap-tiap variabel penelitian:

- a. Persentase balita gizi buruk di Indonesia menurut Provinsi Tahun 2018

Nilai persentase balita gizi buruk di Indonesia diperoleh dari jumlah balita yang dinyatakan memiliki status gizi buruk yang didasarkan pada indeks berat badan menurut umur 0-59 bulan terhadap seluruh jumlah balita umur 0-59 bulan pada tahun 2018 pada setiap provinsi (Kementerian kesehatan Republik Indonesia, 2019).

- b. Persentase bayi mendapat ASI eksklusif di Indonesia menurut Provinsi Tahun 2018

Air Susu Ibu (ASI) eksklusif berdasarkan Peraturan Pemerintah Nomor 33 Tahun 2012 tentang pemberian ASI eksklusif adalah ASI yang diberikan kepada bayi sejak dilahirkan selama enam bulan, tanpa menambahkan dan/atau mengganti makanan atau minuman lain (kecuali obat, vitamin, dan mineral). Salah satu tujuan dari pemberian ASI eksklusif adalah mencegah dan menangani masalah balita gizi buruk. Rumus persentase bayi mendapat ASI eksklusif adalah jumlah bayi yang mendapat hanya ASI saja dibagi dengan jumlah seluruh bayi usia 0-59 bulan dikali 100% (Kementerian kesehatan Republik Indonesia, 2019).

- c. Persentase penduduk miskin di Indonesia menurut Provinsi Tahun 2018

Kemiskinan didefinisikan sebagai kondisi dimana seseorang atau sekelompok orang tidak mampu memenuhi hak-hak dasarnya untuk mempertahankan dan mengembangkan kehidupan yang bermartabat. Kemiskinan juga dipahami sebagai ketidakmampuan ekonomi penduduk untuk memenuhi kebutuhan dasar makanan maupun non makanan yang diukur dari pengeluaran. Persentase penduduk miskin diperoleh dari jumlah penduduk miskin dibagi dengan jumlah penduduk menurut Provinsi di Indonesia (Kementerian kesehatan Republik Indonesia, 2019).

- d. Persentase kabupaten/kota yang melaksanakan PHBS di Indonesia menurut Provinsi Tahun 2018

Perilaku Hidup Bersih dan Sehat (PHBS) merupakan sekumpulan perilaku yang dipraktikkan atas dasar keadaran individu untuk mencegah penyakit atau untuk menjaga kesehatan. Kebijakan PHBS menjadi komponen penting suatu daerah sebagai indikator suatu kebersihan daerah untuk menurunkan kejadian penyakit yang disebabkan oleh perilaku yang tidak sehat. Rumus Persentase kabupaten/kota yang

melaksanakan PHBS adalah jumlah kabupaten/kota yang melaksanakan PHBS dibagi jumlah kabupaten/kota dikali 100% (Kementerian kesehatan Republik Indonesia, 2019).

- e. Persentase posyandu aktif di Indonesia menurut Provinsi Tahun 2018

Posyandu (Pos Pelayanan Terpadu) merupakan salah satu bentuk Upaya Kesehatan Bersumberdaya Masyarakat (UKBM) yang dilaksanakan untuk memberdayakan dan memberikan kemudahan kepada masyarakat guna memperoleh pelayanan kesehatan bagi ibu, bayi, dan anak balita. Posyandu aktif merupakan posyandu yang mampu melaksanakan kegiatan utamanya secara rutin setiap bulan. Rumus persentase posyandu aktif adalah posyandu aktif dibagi total posyandu aktif dikali 100% (Kementerian kesehatan Republik Indonesia, 2019).

- f. Persentase cakupan kunjungan neonatal pertama di Indonesia menurut Provinsi Tahun 2018

Cakupan Kunjungan Neonatal Pertama atau KN1 merupakan indikator yang menggambarkan upaya kesehatan yang dilakukan untuk mengurangi risiko kematian pada periode neonatal pertama yaitu 6-48 jam setelah lahir, sehingga kunjungan neonatal pertama merupakan upaya untuk mencegah balita gizi buruk. Rumus persentase cakupan KN1 adalah cakupan KN1 dibagi dengan jumlah lahir hidup dikali 100% (Kementerian kesehatan Republik Indonesia, 2019).

Dari ke lima variabel prediktor tersebut dilakukan pemilihan variabel untuk dipakai dalam penelitian persentase balita gizi buruk. Variabel-variabel prediktor yang terpilih diduga berpengaruh terhadap variabel respon adalah persentase penduduk miskin dan persentase posyandu aktif di Indonesia tahun 2018.

3.4 Struktur Data

Struktur data yang digunakan dalam penelitian ini ditunjukkan pada Tabel 3.2. Struktur data yang digunakan sesuai dengan jumlah variabel yang digunakan, yaitu satu variabel respon dan lima

variabel prediktor. Sedangkan jumlah provinsi di Indonesia sebanyak 34 Provinsi.

Tabel 3.2 Struktur Data Penelitian

Provinsi	y	x₁	x₂	x₃	x₄	x₅
1	y_1	$x_{1,1}$	$x_{2,1}$	$x_{3,1}$	$x_{4,1}$	$x_{5,1}$
2	y_2	$x_{1,2}$	$x_{2,2}$	$x_{3,2}$	$x_{4,2}$	$x_{5,2}$
3	y_3	$x_{1,3}$	$x_{2,3}$	$x_{3,3}$	$x_{4,3}$	$x_{5,3}$
...
34	y_{34}	$x_{1,34}$	$x_{2,34}$	$x_{3,34}$	$x_{4,34}$	$x_{5,34}$

3.5 Langkah Penelitian

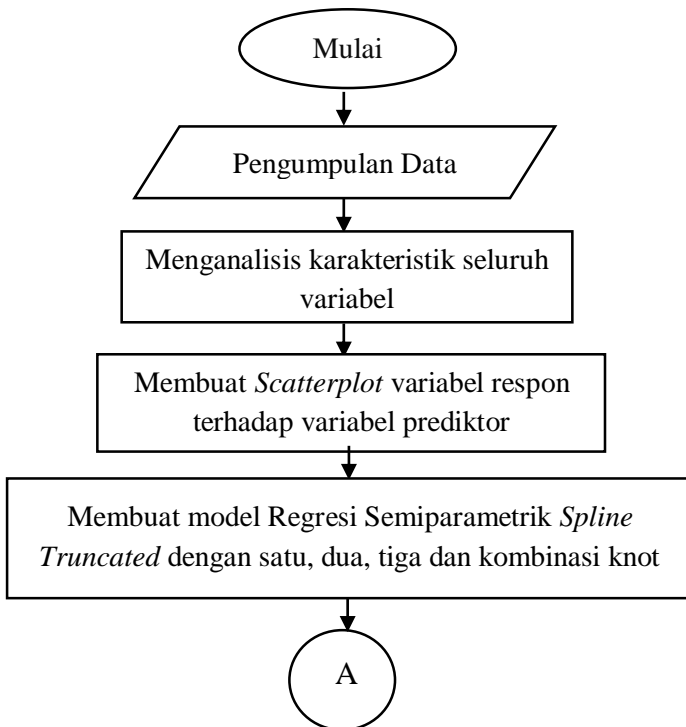
Metode yang digunakan pada penelitian ini adalah Regresi Semiparametrik *Spline Truncated*. Berikut adalah langkah-langkah analisis yang dilakukan dalam penelitian ini sebagai berikut

1. Menganalisis karakteristik presentase balita gizi buruk setiap provinsi beserta faktor-faktor yang diduga mempengaruhinya.
2. Membuat *scatterplot* antara variabel respon yaitu presentase balita gizi buruk setiap provinsi di Indonesia dengan masing-masing variabel prediktor yaitu variabel yang diduga mempengaruhi presentase balita gizi buruk untuk mengetahui pola yang terbentuk.
3. Memodelkan variabel presentase balita gizi buruk setiap provinsi menggunakan model Regresi Semiparametrik *Spline Truncated* dengan satu, dua, dan tiga knot.
4. Menentukan titik knot optimal berdasarkan nilai GCV yang paling minimum.
5. Mendapatkan model Regresi Semiparametrik *Spline Truncated* terbaik dengan titik knot optimal.
6. Melakukan pengujian signifikansi pada parameter Regresi Semiparametrik *Spline Truncated* secara serentak dan parsial.

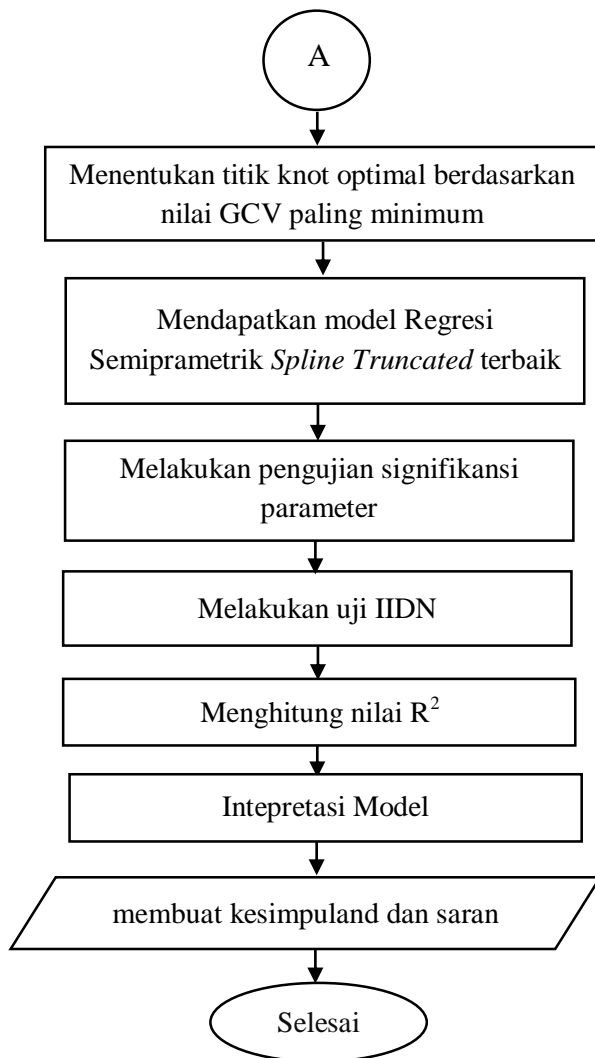
7. Melakukan uji asumsi residual identik, independen, dan berdistribusi normal (IIDN) dari model Regresi Semi-parametrik *Spline Truncated*.
8. Menghitung nilai R^2 .
9. Membuat interpretasi model.
10. Menarik kesimpulan dan saran.

3.6 Diagram Alir

Diagram alir dari langkah penelitian yang dilakukan ditunjukkan pada Gambar 3.2.



Gambar 3.2 Diagram Alir



Gambar 3.2 Diagram Alir (Lanjutan)

(Halaman ini sengaja dikosongkan)

BAB IV

ANALISIS DAN PEMBAHASAN

Bab ini akan membahas mengenai hasil analisis untuk menjawab rumusan masalah pada penelitian ini. Hal-hal yang akan dibahas meliputi karakteristik data persentase balita gizi buruk di Indonesia tahun 2018, mengidentifikasi pola hubungan antara variabel balita gizi buruk dengan masing-masing faktor yang diduga mempengaruhinya menggunakan *scatterplot*, hingga dilakukan pemodelan dengan regresi semiparametrik *spline truncated* menggunakan fungsi *spline* satu knot, dua knot, tiga knot, dan kombinasi knot. Faktor-faktor yang diduga mempengaruhi persentase balita gizi buruk diantaranya adalah persentase penduduk miskin dan persentase posyandu aktif.

4.1 Karakteristik Persentase Balita Gizi Buruk dan Faktor-Faktor yang Diduga Mempengaruhinya

Karakteristik pada data persentase balita gizi buruk dan faktor-faktor yang mempengaruhi balita gizi buruk di Indonesia.

Tabel 4.1 Karakteristik Data

Variabel	<i>Mean</i>	Varians	Min	Maks
Persentase balita gizi buruk (Y)	4,441	2,319	2,000	7,400
Persentase bayi mendapat ASI Eksklusif (X ₁)	61,53	179,40	20,43	80,28
Persentase penduduk miskin (X ₂)	10,608	32,529	3,550	27,430
Persentase yang melakukan PHBS (X ₃)	74,64	878,22	6,90	100,00
Persentase posyandu aktif (X ₄)	55,35	423,93	7,92	99,14
Persentase kunjungan neonatal pertama (X ₅)	89,17	204,68	53,37	128,93

Berdasarkan Tabel 4.1 persentase balita gizi buruk di Indonesia tahun 2018 memiliki rata-rata sebanyak 4,441% dengan varians 2,319%. Persentase balita gizi buruk terkecil mencapai 2% yang terdapat di Provinsi Bali. Sementara persentase balita gizi buruk tertinggi terdapat di Provinsi Maluku mencapai 7,4%. Provinsi dengan persentase gizi buruk terbesar Provinsi Maluku

harus lebih meningkatkan penanggulangan balita gizi buruk agar persentase balita gizi buruk dapat berkurang.

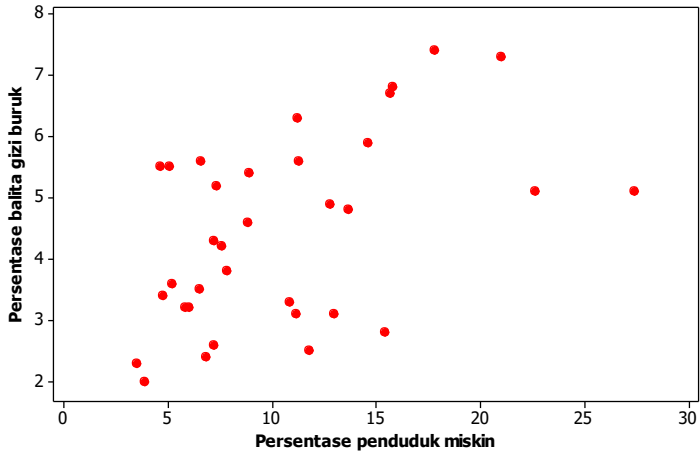
Berikut merupakan karakteristik untuk mendiskripsikan faktor-faktor persentase balita gizi buruk di Indonesia tahun 2018 yang dijelaskan pada variabel prediktor yaitu persentase bayi mendapat ASI Eksklusif, persentase penduduk miskin, persentase kabupaten/kota yang melakukan PHBS, persentase posyandu aktif, dan persentase cakupan kunjungan neonatal pertama.

- a. Persentase bayi mendapat ASI eksklusif dengan karakteristik yang meliputi rata-rata persentase bayi mendapat ASI eksklusif pada tahun 2018 sebesar 61,53% dengan varians 179,40%. Persentase bayi mendapat ASI eksklusif yang terendah terdapat di Provinsi Papua Barat mencapai 20,43% dan tertinggi mencapai 80,28% di Provinsi Sulawesi Barat. Hal ini menunjukkan bahwa persentase bayi mendapat ASI eksklusif tiap Provinsi di Indonesia terletak antara 20,43%-80,28%.
- b. Persentase penduduk miskin dengan karakteristik yang meliputi rata-rata persentase penduduk miskin pada tahun 2018 sebesar 10,608% dengan varians 32,529%. Persentase penduduk miskin yang terendah terdapat di Provinsi DKI Jakarta mencapai 3,55% dan tertinggi mencapai 27,43% di Provinsi Papua. Hal ini menunjukkan bahwa persentase penduduk miskin tiap Provinsi di Indonesia terletak antara 3,55%-27,43%.
- c. Persentase Kabupaten/Kota yang melakukan PHBS dengan karakteristik yang meliputi rata-rata persentase Kabupaten/Kota yang melakukan PHBS pada tahun 2018 sebesar 74,64% dengan varians 878,22%. Persentase Kabupaten/Kota yang melakukan PHBS yang terendah terdapat di Provinsi Papua mencapai 6,90% dan tertinggi mencapai 100,00% di Provinsi Sulawesi Barat. Hal ini menunjukkan bahwa persentase Kabupaten/Kota yang melakukan PHBS tiap Provinsi di Indonesia terletak antara 6,90%-100,00%.

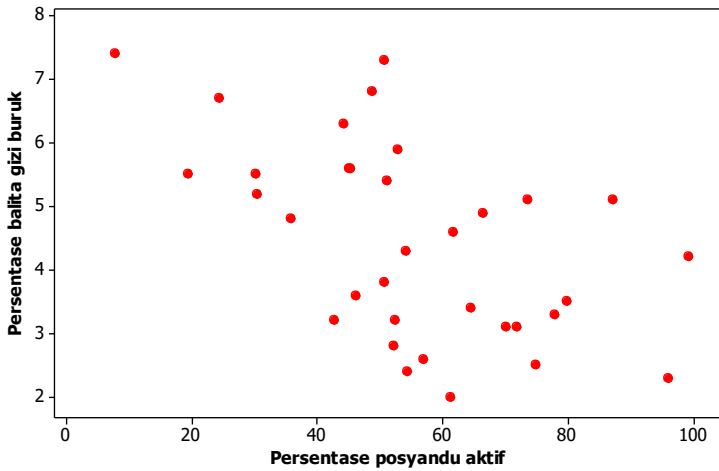
- d. Persentase posyandu aktif dengan karakteristik yang meliputi rata-rata persentase posyandu aktif pada tahun 2018 sebesar 55,35% dengan varians 423,93%. Persentase posyandu aktif yang terendah terdapat di Provinsi Maluku mencapai 7,92% dan tertinggi mencapai 99,14% di Provinsi Sulawesi Utara. Hal ini menunjukkan bahwa persentase posyandu aktif tiap Provinsi di Indonesia terletak antara 7,92%-99,14%.
- e. Persentase cakupan kunjungan neonatal pertama dengan karakteristik yang meliputi rata-rata persentase cakupan kunjungan neonatal pertama pada tahun 2018 sebesar 89,17% dengan varians 204,68%. Persentase cakupan kunjungan neonatal pertama yang terendah terdapat di Provinsi Papua mencapai 53,37% dan tertinggi mencapai 128,93% di Provinsi Jawa Tengah. Hal ini menunjukkan bahwa persentase cakupan kunjungan neonatal pertama tiap Provinsi di Indonesia terletak antara 53,37%-128,93%.

4.2 Pola Hubungan Persentase Balita Gizi Buruk dengan Faktor yang Diduga Mempengaruhi

Pola hubungan antara variabel respon (persentase balita gizi buruk) dengan masing-masing variabel prediktor yang diduga mempengaruhinya dapat menggunakan *scatterplot* (Lampiran 2). Berdasarkan pola hubungan yang ditunjukkan oleh *scatterplot* dapat ditentukan komponen parametrik dan nonparametrik. Sehingga jika telah diketahui komponen-komponen parametrik dan nonparametrik, dapat ditentukan metode yang tepat untuk melakukan pemodelan. Berikut merupakan gambar *scatterplot* variabel yang terpilih yaitu variabel persentase penduduk miskin (X_2) dan variabel posyandu aktif (X_4)



Gambar 4.1 Scatterplot Persentase Balita Gizi Buruk dan Persentase Penduduk Miskin



Gambar 4.2 Scatterplot Persentase Balita Gizi Buruk dan Persentase Posyandu Aktif

Gambar 4.1 dan Gambar 4.2 memberikan ilustrasi hubungan antara variabel persentase balita gizi buruk (Y) dan variabel persentase penduduk miskin (X_2) memiliki kecenderungan perubahan pola perilaku data di interval-interval tertentu, sehingga pendekatan parametrik kurang sesuai diterapkan pada kondisi ini. Metode yang sesuai digunakan untuk menganalisis data dengan kondisi seperti ini adalah *spline truncated*. Dilain pihak, Gambar 4.2 yaitu *scatter plot* hubungan antara variabel persentase balita gizi buruk (Y) dan variabel persentase posyandu aktif (X_4) ada kecenderungan memiliki pola linier, sehingga pendekatan parametrik sesuai digunakan. Dengan menggabungkan pendekatan *spline truncated* (pendekatan nonparametrik) dan pendekatan parametrik didapatkan model pendekatan semiparametrik.

4.3 Pemodelan Persentase Balita Gizi Buruk dengan Regresi Semiparametrik *Spline Truncated*

Dilakukan analisis deskriptif untuk mengetahui karakteristik variabel persentase balita gizi buruk dan masing-masing variabel yang diduga mempengaruhinya. Kemudian dianalisis pola hubungan antar variabel, yang menunjukkan bahwa variabel persentase penduduk miskin (x_2) merupakan komponen non-parametrik dan variabel persentase posyandu aktif (x_4) merupakan komponen parametrik. Sehingga metode yang digunakan dalam penelitian ini adalah dengan menggunakan metode regresi semiparametrik *spline truncated*.

Pemodelan persentase balita gizi buruk menggunakan metode Regresi Semiparametrik *Spline Truncated* dengan satu, dua, dan tiga knot. Pemilihan model terbaik didasari oleh nilai GCV yang paling minimum. Berikut merupakan pemodelan Regresi Semiparametrik *Spline Truncated* dengan satu, dua, dan tiga knot.

4.3.1 Pemilihan Titik Knot Optimum Spline

Pada penelitian ini analisis regresi semiparametrik spline digunakan memodelkan data faktor-faktor yang mempengaruhi balita gizi buruk di Indonesia. Sebelum melakukan pemodelan, hal yang harus dilakukan pertama adalah mencari titik-titik knot

optimum menggunakan metode GCV. Knot yang digunakan dalam penelitian ini yaitu spline truncated linier, kuadratik, dan kubik dengan satu knot, dua knot, dan tiga knot. Dari semua knot tersebut yang akan digunakan untuk membentuk model adalah knot dengan nilai GCV yang paling minimum. Sehingga nantinya akan diperoleh model yang paling baik untuk mengestimasi balita gizi buruk di Indonesia. Berikut adalah analisis menggunakan satu knot, dua knot, dan tiga knot.

a. Regresi Semiparametrik *Spline Truncated Linier*

Pemodelan persentase balita gizi buruk di Indonesia yang diperoleh dari metode Regresi Semiparametrik *Spline Truncated Linier* dengan satu titik knot adalah sebagai berikut.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 (x_{i2} - K_1)_+^1 + \varepsilon_i$$

Menentukan titik knot yang optimal menggunakan satu knot dilakukan 48 kali iterasi. Berikut merupakan lima titik knot pada tiap variabel dengan nilai GCV yang paling minimum.

Tabel 4.2 Nilai GCV Regresi Semiparametrik *Spline Linier* Satu Titik Knot

No	Knot 1	GCV
1	22,55653	1,25202*
2	25,96796	1,25205
3	23,53122	1,25205
4	23,04388	1,25205
5	24,01857	1,25205

*Model spline terbaik

Dari 48 iterasi, Tabel 4.2 menunjukkan bahwa nilai GCV paling minimum menggunakan satu titik knot terdapat pada iterasi ke-39, sama dengan 1,252026. Titik knot optimal untuk variabel persentase penduduk miskin (X_2) berada pada titik knot 22,55653.

Berikut merupakan pemodelan regresi semiparametrik *Spline Truncated Linier* menggunakan satu titik knot dengan GCV yang paling minimum.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 (x_{i2} - 22,55653)_+^1$$

Berikut merupakan pemodelan persentase balita gizi buruk di Indonesia menggunakan Regresi Semiparametrik *Spline Truncated Linier* menggunakan dua titik knot.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 (x_{i2} - K_1)_+^1 + \beta_4 (x_{i2} - K_2)_+^1 + \varepsilon_i$$

Dilakukan sebanyak 1225 kali iterasi untuk menentukan titik knot optimal menggunakan dua titik knot. Berikut merupakan lima titik knot pada tiap variabel dengan nilai GCV yang paling minimum untuk model regresi semiparametrik *Spline Truncated Linier* dengan menggunakan dua titik knot.

Tabel 4.3 Nilai GCV Regresi Semiparametrik *Spline Linier* Dua Titik Knot

No	Knot 1	Knot 2	GCV
1	3,55000	5,98673	1,21333*
2	4,03734	5,98673	1,21333
3	4,52469	5,98673	1,21333
4	5,98673	6,47408	1,21333
5	3,55000	4,52469	1,21333

* Model spline terbaik

Tabel 4.3 menunjukkan bahwa dari 1225 kali iterasi, nilai GCV yang paling minimum menggunakan dua titik knot terdapat pada iterasi ke-5, sama dengan 1,21333. Titik knot paling optimal untuk variabel persentase penduduk miskin adalah $k_1=3,55$ dan $k_2=5,98735$, sehingga diperoleh pemodelan regresi semiparametrik *Spline Truncated Linier* dengan menggunakan dua titik knot sebagai berikut.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 (x_{i2} - 3,55)_+^1 + \hat{\beta}_4 (x_{i2} - 5,986735)_+^1$$

Berikut merupakan pemodelan persentase balita gizi buruk di Indonesia menggunakan Regresi Semiparametrik *Spline Truncated Linier* menggunakan tiga titik knot.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 (x_{i2} - K_1)_+^1 + \beta_4 (x_{i2} - K_2)_+^1 + \beta_5 (x_{i2} - K_3)_+^1 + \varepsilon_i$$

Dilakukan sebanyak 17296 kali iterasi untuk menentukan titik knot optimal menggunakan tiga titik knot. Berikut merupakan lima titik knot pada tiap variabel dengan nilai GCV yang paling

minimum untuk model regresi semiparametrik *Spline Truncated Linier* dengan menggunakan tiga titik knot.

Tabel 4.4 Nilai GCV Regresi Semiparametrik *Spline Linier* Tiga Titik Knot

No	Knot 1	Knot 2	Knot 3	GCV
1	14,75898	15,24633	15,73367	1,09739*
2	14,75898	15,24633	16,22102	1,24060
3	24,50592	25,48061	26,94265	1,25205
4	23,04388	24,50592	24,99327	1,25205
5	23,04388	24,50592	26,94265	1,25205

* Model spline terbaik

Berdasarkan nilai GCV yang paling minimum yang ditunjukkan oleh Tabel 4.4 menggunakan tiga titik knot, diperoleh titik knot yang paling optimal pada iterasi ke-14697, yaitu sama dengan 1,097398. Titik knot optimal untuk variabel persentase penduduk miskin (X_2) adalah $k_1=14,75898$ $k_2=15,24633$, dan $k_3=15,73367$.

Berikut merupakan pemodelan regresi nonparametrik *Spline Truncated Linier* menggunakan tiga titik knot dengan GCV yang paling minimum.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 (x_{i2} - 14,75898)_+^1 + \hat{\beta}_4 (x_{i2} - 15,24633)_+^1 + \hat{\beta}_5 x_{i2} (x_{i2} - 15,73367)_+^1$$

b. Regresi Semiparametrik *Spline Truncated Kuadratik*

Pemodelan persentase balita gizi buruk di Indonesia yang diperoleh dari metode Regresi Semiparametrik *Spline Truncated Kuadratik* dengan satu titik knot adalah sebagai berikut.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 (x_{i2} - K_1)_+^2 + \varepsilon_i$$

Menentukan titik knot yang optimal menggunakan satu knot dilakukan 48 kali iterasi. Berikut merupakan lima titik knot pada tiap variabel dengan nilai GCV yang paling minimum.

Tabel 4.5 Nilai GCV Regresi Semiparametrik *Spline* Kuadratik Satu Titik Knot

No	Knot 1	GCV
1	20,60714	1,32499*
2	20,11980	1,32531
3	21,09449	1,32582
4	19,63245	1,32651
5	21,58184	1,32710

* Model spline terbaik

Dari 48 iterasi, Tabel 4.5 menunjukkan bahwa nilai GCV paling minimum menggunakan satu titik knot terdapat pada iterasi ke-35, sama dengan 1,324997. Titik knot optimal untuk variabel persentase penduduk miskin (X_2) berada pada titik knot 20,60714.

Berikut merupakan pemodelan regresi semiparametrik *Spline Truncated Kuadratik* menggunakan satu titik knot dengan GCV yang paling minimum.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i2}^2 + \hat{\beta}_4 (x_{i2} - 20,60714)_+^2$$

Berikut merupakan pemodelan persentase balita gizi buruk di Indonesia menggunakan Regresi Semiparametrik *Spline Truncated Kuadratik* menggunakan dua titik knot.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 (x_{i2} - K_1)_+^2 + \beta_5 (x_{i2} - K_2)_+^2 + \varepsilon_i$$

Dilakukan sebanyak 1225 kali iterasi untuk menentukan titik knot optimal menggunakan dua titik knot. Berikut merupakan lima titik knot pada tiap variabel dengan nilai GCV yang paling minimum untuk model regresi semiparametrik *Spline Truncated Kuadratik* dengan menggunakan dua titik knot.

Tabel 4.6 Nilai GCV Regresi Semiparametrik *Spline* Kuadratik Dua Titik Knot

No	Knot 1	Knot 2	GCV
1	4,03734	14,27163	1,20622*
2	5,49938	13,78429	1,20624
3	3,55000	14,27163	1,20674
4	5,01204	13,78429	1,20708
5	6,96142	12,80959	1,20737

* Model spline terbaik

Tabel 4.6 menunjukkan bahwa dari 1225 kali iterasi, nilai GCV yang paling minimum menggunakan dua titik knot terdapat pada iterasi ke-70, sama dengan 1.206224. Titik knot paling optimal untuk variabel persentase penduduk miskin adalah $k_1=4,037347$ dan $k_2=14,27163$, sehingga diperoleh pemodelan regresi semiparametrik *Spline Truncated Kuadratik* dengan menggunakan dua titik knot sebagai berikut.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i2}^2 + \hat{\beta}_4 (x_{i2} - 4,037)_+^2 + \hat{\beta}_5 (x_{i2} - 14,271)_+^2$$

Berikut merupakan pemodelan persentase balita gizi buruk di Indonesia menggunakan Regresi Semiparametrik *Spline Truncated Kuadratik* menggunakan tiga titik knot.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 (x_{i2} - K_1)_+^2 + \beta_5 (x_{i2} - K_2)_+^2 + \beta_6 (x_{i2} - K_3)_+^2 + \varepsilon_i$$

Dilakukan sebanyak 17296 kali iterasi untuk menentukan titik knot optimal menggunakan tiga titik knot. Berikut merupakan lima titik knot pada tiap variabel dengan nilai GCV yang paling minimum untuk model regresi semiparametrik *Spline Truncated Kuadratik* dengan menggunakan tiga titik knot.

Tabel 4.7 Nilai GCV Regresi Semiparametrik *Spline Kuadratik* Tiga Titik Knot

No	Knot 1	Knot 2	Knot 3	GCV
1	22,06918	26,45531	26,94265	1,32816*
2	22,55653	26,45531	26,94265	1,32872
3	22,55653	25,96796	26,94265	1,32872
4	22,55653	25,96796	26,45531	1,32872
5	22,55653	25,48061	26,94265	1,32872

* Model spline terbaik

Berdasarkan nilai GCV yang paling minimum yang ditunjukkan oleh Tabel 4.7 menggunakan tiga titik knot, diperoleh titik knot yang paling optimal pada iterasi ke-17176, yaitu sama dengan 1,32. Titik knot optimal untuk variabel persentase penduduk miskin (X_2) adalah $k_1=14,75898$ $k_2=15,24633$, dan $k_3=15,73367$.

Berikut merupakan pemodelan regresi nonparametrik *Spline Truncated* menggunakan tiga titik knot dengan GCV yang paling minimum.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 (x_{i2} - 14,75898)_+^2 + \hat{\beta}_4 (x_{i2} - 15,24633)_+^2 + \hat{\beta}_5 x_{i2} (x_{i2} - 15,73367)_+^2$$

c. Regresi Semiparametrik *Spline Truncated Kubik*

Pemodelan persentase balita gizi buruk di Indonesia yang diperoleh dari metode Regresi Semiparametrik *Spline Truncated Kubik* dengan satu titik knot adalah sebagai berikut.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 x_{i2}^3 + \beta_5 (x_{i2} - K_1)_+^3 + \varepsilon_i$$

Menentukan titik knot yang optimal menggunakan satu knot dilakukan 48 kali iterasi. Berikut merupakan lima titik knot pada tiap variabel dengan nilai GCV yang paling minimum.

Tabel 4.8 Nilai GCV Regresi Semiparametrik *Spline Kubik* Satu Titik Knot

No	Knot 1	GCV
1	8,42346	1,34414*
2	7,93612	1,34783
3	7,44877	1,35092
4	6,96142	1,35338
5	6,47408	1,35558

* Model spline terbaik

Dari 48 iterasi, Tabel 4.8 menunjukkan bahwa nilai GCV paling minimum menggunakan satu titik knot terdapat pada iterasi ke-10, sama dengan 1,344145. Titik knot optimal untuk variabel persentase penduduk miskin (X_2) berada pada titik knot 8,423469.

Berikut merupakan pemodelan regresi semiparametrik *Spline Truncated Kubik* menggunakan satu titik knot dengan GCV yang paling minimum.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i2}^2 + \hat{\beta}_4 (x_{i2} - 1,344145)_+^2$$

Berikut merupakan pemodelan persentase balita gizi buruk di Indonesia menggunakan Regresi Semiparametrik *Spline Truncated Kubik* menggunakan dua titik knot.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 x_{i2}^3 + \beta_5 (x_{i2} - K_1)_+^3 + \beta_6 (x_{i2} - K_2)_+^3 + \varepsilon_i$$

Dilakukan sebanyak 1225 kali iterasi untuk menentukan titik knot optimal menggunakan dua titik knot. Berikut merupakan lima titik knot pada tiap variabel dengan nilai GCV yang paling minimum untuk model regresi semiparametrik *Spline Truncated Kubik* dengan menggunakan dua titik knot.

Tabel 4.9 Nilai GCV Regresi Semiparametrik *Spline Kubik* Dua Titik Knot

No	Knot 1	Knot 2	GCV
1	3,55000	4,03734	1,34626*
2	3,55000	4,52469	1,34852
3	4,03734	4,52469	1,35075
4	3,55000	5,01204	1,35094
5	4,03734	5,01204	1,35324

* Model spline terbaik

Tabel 4.9 menunjukkan bahwa dari 1225 kali iterasi, nilai GCV yang paling minimum menggunakan dua titik knot terdapat pada iterasi ke-1, sama dengan 1,34626. Titik knot paling optimal untuk variabel persentase penduduk miskin adalah $k_1=3,55$ dan $k_2=4,037347$, sehingga diperoleh pemodelan regresi semiparametrik *Spline Truncated Kubik* dengan menggunakan dua titik knot sebagai berikut.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i2}^2 + \hat{\beta}_4 x_{i2}^3 + \hat{\beta}_5 (x_{i2} - 3,55)_+^3 + \hat{\beta}_6 (x_{i2} - 4,037347)_+^3$$

Berikut merupakan pemodelan persentase balita gizi buruk di Indonesia menggunakan Regresi Semiparametrik *Spline Truncated Kubik* menggunakan tiga titik knot.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 x_{i2}^3 + \beta_5 (x_{i2} - K_1)_+^3 + \beta_6 (x_{i2} - K_2)_+^3 + \beta_7 (x_{i2} - K_3)_+^3 + \varepsilon_i$$

Dilakukan sebanyak 17296 kali iterasi untuk menentukan titik knot optimal menggunakan tiga titik knot. Berikut merupakan

lima titik knot pada tiap variabel dengan nilai GCV yang paling minimum untuk model regresi semiparametrik *Spline Truncated Kubik* dengan menggunakan tiga titik knot.

Tabel 4.10 Nilai GCV Regresi Semiparametrik *Spline Kubik* Tiga Titik Knot

No	Knot 1	Knot 2	Knot 3	GCV
1	4,03734	5,49938	26,94265	1,35795*
2	4,03734	5,01204	26,94265	1,35965
3	4,03734	4,52469	26,94265	1,36524
4	15,24633	26,45531	26,94265	1,38104
5	14,75898	26,45531	26,94265	1,38122

* Model spline terbaik

Berdasarkan nilai GCV yang paling minimum yang ditunjukkan oleh Tabel 4.10 menggunakan tiga titik knot, diperoleh titik knot yang paling optimal pada iterasi ke-135, yaitu sama dengan 1,357956. Titik knot optimal untuk variabel persentase penduduk miskin (X_2) adalah $k_1=4,03734$, $k_2=5,499388$, dan $k_3=26,94265$. Berikut merupakan pemodelan regresi nonparametrik *Spline Truncated Linier* menggunakan tiga titik knot dengan GCV yang paling minimum.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i2}^2 + \hat{\beta}_4 x_{i2}^3 + \hat{\beta}_5 (x_{i2} - 4,037347)_+^3 + \hat{\beta}_6 (x_{i2} - 5,499388)_+^3 + \hat{\beta}_7 (x_{i2} - 26,94265)_+^3$$

4.3.2 Pemilihan Model Terbaik

Pemilihan model terbaik didasari pada titik knot optimal yang ditunjukkan oleh nilai GCV yang paling minimum. Berdasarkan pemodelan yang telah dilakukan dengan berbagai pendekatan linier, kuadratik, dan kubik masing-masing pendekatan terdapat titik knot (satu, dua, dan tiga), kemudian masing-masing nilai GCV pada tiap titik knot yang paling optimal dari masing-masing pendekatan tersebut dibandingkan. Berikut nilai GCV minimum pada tiap pemodelan menggunakan satu, dua, tiga dan kombinasi knot.

Tabel 4.11 Perbandingan Nilai GCV Minimum tiap Titik Knot

	Knot	GCV
Linier	Tiga Knot	1,09739*
Kuadratik	Dua Knot	1,20622
Kubik	Satu Knot	1,34414

*Model Terbaik

Nilai GCV yang paling minimum pada tiap titik knot, menunjukkan bahwa pemodelan semiparametrik *spline truncated* menggunakan pendekatan linier dengan tiga titik knot merupakan model terbaik karena menghasilkan GCV yang paling optimal karena.

4.4 Penaksir Parameter Model Regresi Semiparametrik *Spline Truncated Linier*

Berdasarkan nilai GCV yang paling minimum titik knot paling optimal yang digunakan untuk memodelkan data persentase balita gizi buruk di Indonesia tahun 2018 adalah semiparametrik *spline truncated linier* dengan tiga titik knot, sehingga dapat diestimasi parameter untuk model regresi semiparametrik *spline truncated*. Estimasi parameter menggunakan metode *Ordinary Least Square* (OLS) adalah sebagai berikut

$$\hat{y} = 4,938 - 0,035x_4 + 0,138x_2 - 8,848(x_2 - 14,759)_+^1 + 20,288(x_2 - 15,246)_+^1 + \\ -11,608(x_2 - 15,734)_+^1$$

4.5 Pengujian Signifikan

Pengujian signifikansi parameter model dilakukan untuk mengetahui apakah variabel prediktor yang diduga berpengaruh signifikan terhadap persentase balita gizi buruk di Indonesia atau tidak. Pengujian parameter model regresi semiparametrik *spline truncated* terdiri dari pengujian secara serentak dan pengujian secara parsial atau individu sebagai berikut.

4.5.1 Uji Serentak

Uji parameter secara serentak dilakukan untuk melihat apakah seluruh parameter dari variabel prediktor yang digunakan

berpengaruh signifikan terhadap model. Hipotesis yang digunakan untuk uji parameter secara serentak adalah sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_6 = 0$$

$$H_1 : \text{Minimal terdapat satu } \beta_j \neq 0, j = 1, 2, \dots, 6$$

Berikut hasil uji parameter secara serentak ditampilkan pada Tabel 4.12.

Tabel 4.12 ANOVA Model Regresi Semiparametrik *Spline*

Sumber Variasi	df	SS	MS	F	P-Value
Regresi	5	51,23765	10,24753	11,33903	$4,79 \times 10^{-6}$
Error	28	25,30471	0,9037395		
Total	33	76,54235			

Berdasarkan Tabel 4.12 dapat dilihat bahwa F-hitung sebesar 11,33903 dan nilai p-value sebesar 0,000004 maka dapat diambil keputusan Tolak H_0 apabila nilai statistik uji $F_{hitung} > F_{(0,05;5;28)}$ sebesar $11,33903 > 2,56$ maka menghasilkan keputusan tolak H_0 . Sehingga dapat disimpulkan bahwa minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap persentase balita gizi buruk di Indonesia. Untuk mengetahui variabel mana yang berpengaruh maka dilakukan uji parsial.

4.5.2 Uji Parsial

Uji parsial dilakukan untuk mengetahui variabel mana yang memberikan pengaruh signifikan terhadap persentase balita gizi buruk secara individu atau parsial. Berikut hipotesis untuk melakukan uji parameter secara parsial.

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, \dots, 6$$

Tabel 4.13 Hasil Pengujian Signifikansi Parameter Secara Individu

Variabel	Parameter	Estimator	P-value	t_{hitung}	Keputusan
Konstan	β_0	4,938	$1,287 \times 10^{-7}$	7,003	Tolak H_0
x_1	β_1	-0,035	0,00042	-3,911	Tolak H_0

Tabel 4.13 Hasil Pengujian Signifikansi Parameter Secara Individu (Lanjutan)

Variabel	Parameter	Estimator	P-value	<i>t</i> _{hitung}	Keputusan
x_2	β_2	0,138	0,02502	2,368	Tolak H_0
	β_3	-8,848	0,00912	-2,801	Tolak H_0
	β_4	20,288	0,00468	3,073	Tolak H_0
	β_5	-11,608	0,00383	-3,153	Tolak H_0

Berdasarkan Tabel 4.13 diperoleh kesimpulan Tolak H_0 , yaitu semua parameter. Hal ini menunjukkan semua parameter parameter berpengaruh signifikan terhadap model.

4.6 Uji Asumsi Residual

Uji asumsi residual dilakukan untuk mengetahui apakah residual yang dihasilkan dari model regresi tersebut telah memenuhi asumsi yakni identik, independen, dan berdistribusi normal. Hasil pengujian asumsi residual adalah sebagai berikut.

4.6.1 Asumsi Identik

Asumsi residual identik dilakukan untuk mengetahui apakah residual memiliki varians yang telah homogen atau tidak terjadi heterokedastisitas. Apabila asumsi residual identik tidak terpenuhi atau dapat disebut juga dengan terjadinya heterokedastisitas akan mengakibatkan kerugian pada efisiensi estimator. Berikut adalah hipotesis pengujian asumsi identik.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{Minimal ada satu } \sigma_i^2 \neq \sigma^2 ; i = 1, 2, \dots, 34$$

Pengujian asumsi residual identik menggunakan uji *Glejser* Dengan menggunakan taraf kepercayaan 95% maka diperoleh nilai $F_{(0,05;5;28)} = 2,56$. Hasil pengujian asumsi identik ditampilkan pada Tabel 4.14.

Tabel 4.14 Hasil Pengujian Statistik Uji *Glejser*

Sumber	Df	SS	MS	F _{hit}	P-value
Regresi	5	1,691236	0,3382472	1,676948	0,17289
Error	28	5,647713	0,201704		
Total	33	7,33895			

Berdasarkan Tabel 4.14 diperoleh nilai statistik uji F_{hitung} sama dengan 0,8892 dengan $p\text{-value} = 0,5907$. Karena nilai statistik uji $F_{hitung} < F_{(0,05;5;28)}$, maka menghasilkan keputusan gagal tolak H_0 . Selain itu diperoleh $p\text{-value} > 5\%$ dan diperoleh keputusan gagal tolak H_0 . Hal ini menunjukkan tidak terjadi heterokedastisitas pada model, sehingga asumsi residual identik terpenuhi.

4.6.2 Asumsi Independen

Pengujian asumsi residual independen menggunakan *Durbin-Watson* untuk mengetahui residual data model regresi semiparametric spline sudah memenuhi asumsi independen atau tidak. Hipotesis pengujian asumsi residual independen adalah sebagai berikut.

Tabel 4.15 Uji Asumsi Residual Independen

<i>Durbin-Watson statistic</i>	dL	dU
2,092042	1,333	1,580

Tabel 4.15 menunjukkan hasil uji asumsi residual independen dengan menggunakan uji Durbin-Watson diperoleh keputusan bahwa gagal tolak H_0 karena nilai Durbin-Watson 2,092 lebih besar dari nilai dL 1,33 dan lebih kecil juga dari nilai dU sebesar 1,580 sehingga kesimpulannya adalah residual data memenuhi asumsi independen.

4.6.3 Asumsi Distribusi Normal

Pengujian asumsi berdistribusi normal dilakukan untuk mengetahui apakah residual suatu data telah mengikuti distribusi normal atau tidak dengan *mean* nol dan varians σ^2 . Uji yang digunakan adalah uji *Kolmogorov-Smirnov*.

Hipotesis :

$$H_0 : F_n(\varepsilon) = F_0(\varepsilon)$$

$$H_1 : F_n(\varepsilon) \neq F_0(\varepsilon)$$

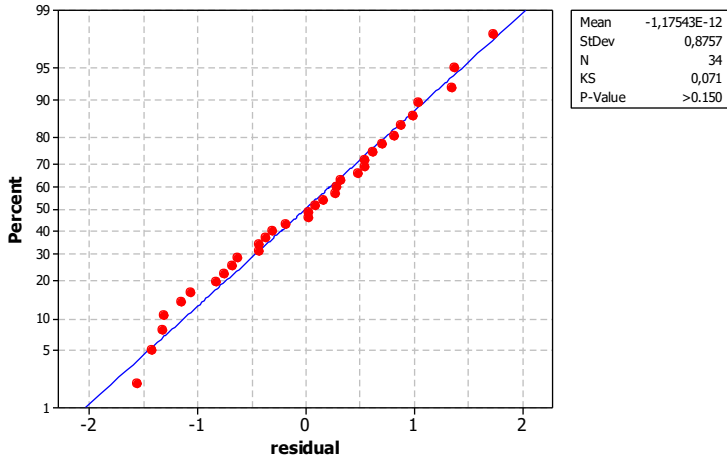
atau

H_0 : residual berdistribusi normal

H_1 : residual tidak berdistribusi normal

Statistik uji :

$$D = \sup_{\varepsilon} |F_n(\varepsilon) - F_0(\varepsilon)|$$



Gambar 4.3 Hasil Uji *Kolmogrov Smirnov* dan Plot Normalitas Residual

Titik plot residual yang ditunjukkan pada Gambar 4.12 menunjukkan penyebaran titik plot residual berada pada sumbu diagonal dari grafik, sehingga hal ini mengindikasikan residual berdistribusi normal. Sehingga dapat ditarik kesimpulan bahwa residual mengikuti pola distribusi normal, maka model telah memenuhi asumsi residual distribusi normal.

4.7 Koefisien Determinasi (R^2)

Nilai koefisien determinasi menunjukkan seberapa baiknya model yang digunakan. Menurut perhitungan nilai koefisien determinasi R^2 diperoleh sebagai berikut.

$$R^2 = \frac{SS_{\text{Regresi}}}{SS_{\text{total}}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 66,94\%$$

Berdasarkan hasil perhitungan menggunakan persamaan koefisien determinasi (R^2) diperoleh sama dengan 66,94%. Artinya, model regresi semiparametrik *spline truncated linier* dengan tiga knot mampu menjelaskan variabilitas persentase balita gizi buruk di Indonesia sama dengan 66,94%.

4.8 Interpretasi Model *Spline* Terbaik pada Variabel Persentase Penduduk Miskin (X_2)

Mengasumsikan semua variabel selain variabel X_2 konstan, dimana X_2 merupakan komponen nonparametrik maka persamaan regresi dari persentase penduduk miskin (X_2) terhadap persentase balita gizi buruk di Indonesia (Y) adalah sebagai berikut.

$$\hat{y} = \begin{cases} 4,94 + 0,14x_1 & ; x_2 < 14,76 \\ 135,53 - 8,70x_1 & ; 14,76 \leq x_2 < 15,25 \\ -173,79 + 11,58x_1 & ; 15,25 \leq x_2 < 15,73 \\ 8,85 - 0,03x_1 & ; x_2 \geq 15,73 \end{cases}$$

Menggunakan tiga titik knot maka diperoleh empat interval, dimana untuk setiap interval memiliki interpretasi model yang berbeda. Pada interval $x_2 < 14,76$ menunjukkan bahwa apabila persentase penduduk miskin di suatu provinsi $< 14,76$ maka setiap kenaikan persentase penduduk miskin sebesar satu satuan akan meningkatkan persentase balita gizi buruk sebesar 0,14%.

Persentase penduduk miskin di suatu provinsi berada pada interval $14,76 \leq x_2 < 15,25$ maka setiap kenaikan persentase penduduk miskin sebesar satu satuan mengakibatkan persentase balita gizi buruk naik sebanyak 11,58%.

Pada interval $x_2 \geq 15,73$ menunjukkan bahwa apabila persentase penduduk miskin sebesar satu satuan mengakibatkan persentase balita gizi buruk turun sebanyak 0,03%.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan analisis yang telah dilakukan, diperoleh kesimpulan sebagai berikut.

1. Rata-rata persentase balita gizi buruk di Indonesia tahun 2018 adalah 4,441%. Persentase balita gizi buruk terkecil mencapai 2% yang terdapat di Provinsi Bali. Sementara persentase balita gizi buruk tertinggi terdapat di Provinsi Maluku mencapai 7,4%. Provinsi dengan persentase gizi buruk terbesar Provinsi Maluku harus lebih meningkatkan penanggulangan balita gizi buruk agar persentase balita gizi buruk dapat berkurang. persentase bayi mendapat ASI eksklusif tiap Provinsi di Indonesia terletak antara 20,43%-80,28%. Variabel persentase penduduk miskin tiap Provinsi di Indonesia terletak antara 3,55%-27,43%. Variabel persentase Kabupaten/Kota yang melakukan PHBS tiap Provinsi di Indonesia terletak antara 6,90%-100,00%. Variabel persentase posyandu aktif tiap Provinsi di Indonesia terletak antara 7,92%-99,14%. Variabel persentase cakupan kunjungan neonatal pertama tiap Provinsi di Indonesia terletak antara 53,37%-128,93%.
2. Model regresi terbaik yang digunakan untuk memodelkan persentase balita gizi buruk di Indonesia tahun 2018 adalah model regresi semiparametrik *spline truncated linier* dengan menggunakan tiga titik knot. Model ini memiliki nilai GCV paling minimum yaitu 1,097398 dengan nilai R^2 sebesar 66,94%. Bentuk model regresi semiparametrik *spline truncated linier* dengan tiga titik knot yang adalah sebagai berikut.

$$\hat{y} = 4,938 - 0,035x_4 + 0,138x_2 - 8,848(x_2 - 14,759)_+^1 + 20,288(x_2 - 15,246)_+^1 - 11,608(x_2 - 15,734)_+^1$$

5.2 Saran

Berdasarkan penelitian yang telah dilakukan, saran yang dapat direkomendasikan untuk penelitian selanjutnya adalah menambah variabel lain yang diduga mempengaruhi balita gizi buruk, Kemudian mempertimbangkan pola hubungan antara variabel prediktor dengan variabel respon, untuk dapat menentukan metode yang tepat digunakan untuk memodelkan persentase balita gizi buruk di Indonesia tahun 2018.

DAFTAR PUSTAKA

- Budiantara, I. N. (2001). *Estimasi Parametrik dan Nonparametrik untuk pendekatan Kurva Regresi, Seminar Nasional Statistika V, Jurusan Statistika*. Surabaya: ITS.
- Budiantara, I. N. (2007). *Kelas Estimator Linier Dlam Regresi Parametrik dan Semiparametrik*. Pembicara Utama pada Seminar Nasional Matematika FMIPA Universitas Lambung Mengkurat (UNLAM), Banjarmasin.
- Carrol, R., Wand , M., & Ruppert, D. (2003). *Cambridge Series in Statistical and Probabilistic Mathematics Semiparametric Regression*. United States of America: Cambridge University Press.
- Daniel, W. W. (1989). *Statistika Non Parametrik, Alih Bahasa : Alex Tri Kuncoro*. Jakarta: PT. Gramedia.
- Devi, M. (2010). Analisis Faktor-faktor yang Berpengaruh Terhadap Statu Gizi Balita di Pedesaan. *Teknologi dan Kejuruan*, 183-192.
- Dhantiasih, L. J. (2014). *Cara Mendekteksi Gizi Buruk pada Balita*. Sarihusada Nutrisi untuk Bangsa.
- Draper, N. R., & Smith, H. (1998). *Analisis Regresi Terapan. Alih Bahasa: Ir. Bambang Sumantri*. Jakarta: Gramedia Pustaka Utama.
- Eubank, R. L. (1999). *Nonparametric Regression and Spline Smoothing* (2ND ed.). New York: Marcel Dekker Inc.
- Gujarati, D. (2003). *Basic Econometrics (Ekonometrika Dasar)*. Alih bahasa : Sumarno Zain. Jakarta: Erlangga.

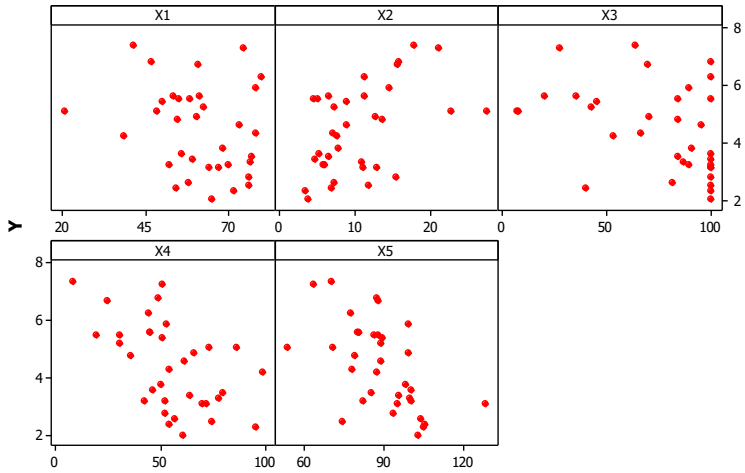
- Gujarati, D. (2004). *Basic Econometrics (4th Edition)*. New York: McGraw-Hill.
- Hardle, W. (1990). *Applied Nonparametric Regression*. New York: Cambridge University Press.
- Kementerian kesehatan Republik Indonesia. (2019). *Profil Kesehatan Indonesia Tahun 2018*. Jakarta: Pusat Data dan Informasi Kementerian Kesehatan Republik Indonesia.
- Kusriadi. (2010). *Analisis Faktor Risiko yang Mempengaruhi Kejadian Kurang Gizi Pada Anak Balita DI Provinsi NTB*. Bogor: IPB.
- Novitasari, D. (2012). Faktor-faktor Risiko Kejadian Gizi Buruk pada Balita yang Dirawat Di RSUP Dr.Kariadi Semarang. *Laporan Akhir Hasil Penelitian Karya Tulis Ilmiah*.
- Rahim, F. K. (2014). Faktor Resiko Underweight Balita Umur 7-59 Bulan. *Kesehatan Masyarakat*, 9, 115-121.
- Saputra, W., & Nurriszka, R. H. (2012). Faktor Demografi dan Resiko Gizi Buruk dan Gizi Kurang. *Makara Kesehatan*, 95-101.
- Wahba, G. (1990). *Spline Models For Observation Data* . Pennsylvania: SIAM.
- Walpole, R. (1993). *Pengantar Metode Statistika, Edisi Ketiga, Alih Bahasa : Bambang Sumantri*. Jakarta: PT Gramedia Pusaka Utama.

LAMPIRAN

Lampiran 1. Data Persentase Balita Gizi Buruk di Indonesia Tahun 2018 dengan Faktor-faktor yang mempengaruhi

Provinsi	Y	X1	X2	X3	X4	X5
Aceh	6.7	60.84	15.68	69.57	24.51	88.2
Sumatra Utara	5.4	50.07	8.94	45.45	51.26	89.67
Sumatra Barat	3.5	77.09	6.55	84.21	79.89	85.48
Riau	4.3	78.33	7.21	66.67	54.19	78.06
Jambi	3.8	68.64	7.85	90.91	50.71	98.29
Sumatra Selatan	4.9	60.43	12.82	70.59	66.37	99.55
Bengkulu	2.8	76.38	15.41	100	52.24	93.86
Lampung	3.1	67.01	13.01	100	70.21	95.39
Kep. Bangka Belitung	3.4	59.11	4.77	100	64.57	95.91
Kepulauan Riau	3.2	52.18	5.83	100	52.53	100.62
DKI Jakarta	2.3	71.88	3.55	100	95.89	105.04
Jawa Barat	2.6	57.97	7.25	81.48	57.08	104.15
Jawa Tengah	3.1	64.19	11.19	100	71.97	128.93
DI Yogyakarta	2.5	76.17	11.81	100	74.89	74.54
Jawa Timur	3.3	76.98	10.85	86.84	77.98	100.21
Banten	3.6	56.1	5.25	100	46.2	100.57
Bali	2	65.04	3.91	100	61.31	102.92
Nusa Tenggara Barat	5.9	78.63	14.63	90	52.82	99.59
Nusa Tenggara Timur	7.3	74.58	21.03	27.27	50.78	63.36
Kalimantan Barat	5.2	62.83	7.37	42.86	30.51	88.96
Kalimantan Tengah	5.5	58.66	5.1	100	19.44	86.46
Kalimantan Selatan	5.5	55.31	4.65	84.62	30.27	87.99
Kalimantan Timur	3.2	70.07	6.06	90	42.76	82.36
Kalimantan Utara	2.4	54.29	6.86	40	54.44	105.83
Sulawesi Utara	4.2	38.69	7.59	53.33	99.14	87.26
Sulawesi Tengah	4.8	54.53	13.69	84.62	35.95	79.32
Sulawesi Selatan	4.6	73.56	8.87	95.83	61.75	88.8
Sulawesi Tenggara	5.6	53.48	11.32	35.29	45.25	80.83
Gorontalo	6.8	46.91	15.83	100	48.91	87.63
Sulawesi Barat	6.3	80.28	11.22	100	44.2	77.7
Maluku	7.4	41.51	17.85	63.64	7.92	70.3
Maluku Utara	5.6	61.37	6.62	20	45.33	80.03
Papua Barat	5.1	20.43	22.66	7.69	87.06	70.65
Papua	5.1	48.32	27.43	6.9	73.67	53.37

Lampiran 2. *Scatter plot* Variabel Y terhadap variabel X_1 , X_2 , X_3 , X_4 , dan X_5



Lampiran 3. *Syntax* Pemilihan Titik Knot Optimal dengan Linier Satu Titik Knot Menggunakan R.

```

library(pracma)
library(MASS)

gcv1=function(para)
{
  data=read.table("F:/semester AKHIR/TA/wanin/dataTA1.txt",header=T)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  dataA=data[, (para+2):q]
  F=matrix(0,nrow=p,ncol=p)
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot1=matrix(ncol=m,nrow=nk)
  for (j in (1:nk))
  {
    a=seq(min(dataA),max(dataA),length.out=50)
    knot1[j]=a[j]
  }
  a1=length(knot1[,1])
  knot1=knot1[2:(a1-1),]
  knot1=data.frame(knot1)
  aa=rep(1,p)
  data1=matrix(ncol=m,nrow=p)
  data2=data[,2:q]
  a2=nrow(knot1)
  GCV=rep(NA,a2)
  Rsq=rep(NA,a2)
  for (i in 1:a2)
  {
    for (j in 1:m)
    {
      for (k in 1:p)
      {
        if(data[k,(j+para+1)]<knot1[i,j]) data1[k,j]=0 else data1[k,j]=data[k,(j+para+1)]-
knot1[i,j]
      }
    }
    mx=cbind(aa,data2,data1)
    mx=as.matrix(mx)
    C=pinv(t(mx)%*%mx)
    B=C%*%(t(mx)%*%data[,1])
    yhat=mx%*%B
    SSE=0
  }
}

```

```

SSR=0
for (r in (1:p))
{
  sum=(data[r,1]-yhat[r,])^2
  sum1=(yhat[r,]-mean(data[,1]))^2
  SSE=SSE+sum
  SSR=SSR+sum1
}
Rsqr[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsqr=as.matrix(Rsqr)
cat("=====", "\n")
cat("Nilai Knot dengan Spline linier 1 knot", "\n")
cat("=====", "\n")
cat("=====", "\n")
cat("Rsqr dengan Spline linier 1 knot", "\n")
cat("=====", "\n")
print (Rsqr)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linier 1 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
print(max(Rsqr))
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linier 1 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.table(GCV, file="F:/semester AKHIR/TA/wanin/output GCV.txt")
write.table(Rsqr, file="F:/semester AKHIR/TA/wanin/output Rsqr.txt")
write.table(knot1, file="F:/semester AKHIR/TA/wanin/output knot.txt")
}
gcv1(1)

```


Lampiran 4. *Syntax* Pemilihan Titik Knot Optimal dengan Linier Dua Titik Knot Menggunakan R.

```

gcv2=function(para)
{
  data=read.table("F:/semester AKHIR/TA/wanin//dataTA1.txt",header=T)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  F=matrix(0,nrow=p,ncol=p)
  dataA=data[, (para+2):q]
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot=matrix(ncol=m,nrow=nk)
  for (j in (1:nk))
  {
    a=seq(min(dataA),max(dataA),length.out=50)
    knot[j]=a[j]
  }
  z=(nk*(nk-1)/2)
  knot2=cbind(rep(NA,(z+1)))
  for (i in (1:m))
  {
    knot1=rbind(rep(NA,2))
    for (j in 1:(nk-1))
    {
      for (k in (j+1):nk)
      {
        xx=cbind(knot[j,i],knot[k,i])
        knot1=rbind(knot1,xx)
      }
    }
    knot2=cbind(knot2,knot1)
  }
  knot2=knot2[2:(z+1),2:(2*m+1)]
  aa=rep(1,p)
  data2=matrix(ncol=(2*m),nrow=p)
  data1=data[,2:q]
  a1=length(knot2[,1])
  GCV=rep(NA,a1)
  Rsq=rep(NA,a1)
  for (i in 1:a1)
  {
    for (j in 1:(2*m))
    {
      if(mod(j,2)==1) b=floor(j/2)+1 else b=j/2
      for (k in 1:p)

```

```

{
  if (data1[k,b]<knot2[i,j]) data2[k,j]=0 else data2[k,j]=data1[k,b]-knot2[i,j]
}
}
print(data1)
mx=cbind(aa,data1,data2)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in (1:p))
{
  sum=(data[r,1]-yhat[r,])^2
  sum1=(yhat[r,]-mean(data[,1]))^2
  SSE=SSE+sum
  SSR=SSR+sum1
}
Rsq[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsq=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
cat("=====", "\n")
cat("Rsq dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
s1=min(GCV)
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
cat(" GCV =",s1, "\n")
write.table(GCV,file="F:/semester AKHIR/TA/wanin//output GCV2.txt")
write.table(Rsq,file="F:/semester AKHIR/TA/wanin//output Rsq2.txt")
write.table(knot2,file="F:/semester AKHIR/TA/wanin//output knot2.txt")
write.table(mx,file="F:/semester AKHIR/TA/wanin//mx.txt")
}
gcv2(1)

```

Lampiran 5. *Syntax* Pemilihan Titik Knot Optimal dengan Linier Tiga Titik Knot Menggunakan R.

```

gcv3=function(para)
{
  data=read.table("F:/semester AKHIR/TA/wanin/dataTA1.txt",header=T)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  F=matrix(0,nrow=p,ncol=p)
  dataA=data[, (para+2):q]
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot=matrix(ncol=m,nrow=nk)
  for (j in (1:nk))
  {
    a=seq(min(dataA),max(dataA),length.out=50)
    knot[j]=a[j]
  }
  knot=knot[2:(nk-1),]
  knot=data.frame(knot)
  a2=nrow(knot)
  z=(a2*(a2-1)*(a2-2))/6
  knot1=cbind(rep(NA,(z+1)))
  for (i in (1:m))
  {
    knot2=rbind(rep(NA,3))
    for (j in 1:(a2-2))
    {
      for (k in (j+1):(a2-1))
      {
        for (g in (k+1):a2)
        {
          xx=cbind(knot[j,i],knot[k,i],knot[g,i])
          knot2=rbind(knot2,xx)
        }
      }
    }
    knot1=cbind(knot1,knot2)
  }
  knot1=knot1[2:(z+1),2:(3*m+1)]
  aa=rep(1,p)
  data1=matrix(ncol=(3*m),nrow=p)
  data2=data[, (para+2):q]
  a1=length(knot1[,1])
  GCV=rep(NA,a1)
  Rsq=rep(NA,a1)
}

```

```

data2=data.frame(data2)
for (i in 1:a1)
{
  for (j in 1:ncol(knot1))
  {
    b=ceiling(j/3)
    for (k in 1:p)
    {
      if (data2[k,b]<knot1[i,j]) data1[k,j]=0 else data1[k,j]=data2[k,b]-knot1[i,j]
    }
  }
  mx=cbind(aa,data[,2:q],data1)
  mx=as.matrix(mx)
  C=pinv(t(mx)%*%mx)
  B=C%*%(t(mx)%*%data[,1])
  yhat=mx%*%B
  SSE=0
  SSR=0
  for (r in (1:p))
  {
    sum=(data[r,1]-yhat[r,])^2
    sum1=(yhat[r,]-mean(data[,1]))^2
    SSE=SSE+sum
    SSR=SSR+sum1
  }
  Rsq[i]=(SSR/(SSE+SSR))*100
  MSE=SSE/p
  A=mx%*%C%*%t(mx)
  A1=(F-A)
  A2=(sum(diag(A1))/p)^2
  GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsq=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
cat("=====", "\n")
cat("Rsq dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (Rsq)
r=max(Rsq)
print (r)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)

```

```
cat("=====","\\n")
cat("HASIL GCV terkecil dengan Spline linear 3 knot","\\n")
cat("=====","\\n")
cat(" GCV =",s1,"\\n")
write.table(GCV,file="F:/semester AKHIR/TA/wanin/output GCV3.txt")
write.table(Rsq,file="F:/semester AKHIR/TA/wanin/output Rsq3.txt")
write.table(knot1,file="F:/semester AKHIR/TA/wanin/output knot3.txt")

}
gcv3(1)
```

Lampiran 6. *Syntax* Pemilihan Titik Knot Optimal dengan Kuadratik Satu Titik Knot Menggunakan R.

```

gcv1=function(para)
{
  data=read.table("F:/semester AKHIR/TA/wanin/dataTA1.txt",header=T)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  dataA=data[, (para+2):q]
  F=matrix(0,nrow=p,ncol=p)
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot1=matrix(ncol=m,nrow=nk)
  for (j in (1:nk))
  {
    a=seq(min(dataA),max(dataA),length.out=50)
    knot1[j]=a[j]
  }
  a1=length(knot1[,1])
  knot1=knot1[2:(a1-1),]
  knot1=data.frame(knot1)
  aa=rep(1,p)
  data1=matrix(ncol=m,nrow=p)
  data2=data[,2:q]
  a2=nrow(knot1)
  GCV=rep(NA,a2)
  Rsq=rep(NA,a2)
  for (i in 1:a2)
  {
    for (j in 1:m)
    {
      for (k in 1:p)
      {
        if(data[k,(j+para+1)]<knot1[i,j]) data1[k,j]=0 else data1[k,j]=data[k,(j+para+1)]-
knot1[i,j]
      }
    }
  }
  mx=cbind(aa,data2,data2[,2]^2,data1^2)
  mx=as.matrix(mx)
  C=pinv(t(mx)%*%mx)
}

```

```

B=C%%(t(mx)%data[,1])
yhat=mx%%B
SSE=0
SSR=0
for (r in (1:p))
{
  sum=(data[r,1]-yhat[r,])^2
  sum1=(yhat[r,]-mean(data[,1]))^2
  SSE=SSE+sum
  SSR=SSR+sum1
}
Rsq[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%%C%%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsq=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline kuadratik 1 knot", "\n")
cat("=====", "\n")
cat("=====", "\n")
cat("Rsq dengan Spline kuadratik 1 knot", "\n")
cat("=====", "\n")
print (Rsq)
cat("=====", "\n")
cat("HASIL GCV dengan Spline kuadratik 1 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
print(max(Rsq))
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline kuadratik 1 knot", "\n")
cat("=====", "\n")
cat(" GCV =",s1, "\n")
write.table(GCV,file="F:/semester AKHIR/TA/wanin/output GCV.txt")
write.table(Rsq,file="F:/semester AKHIR/TA/wanin/output Rsq.txt")
write.table(knot1,file="F:/semester AKHIR/TA/wanin/output knot.txt")
}
gcv1(1)

```

Lampiran 7. *Syntax* Pemilihan Titik Knot Optimal dengan Kuadratik Dua Titik Knot Menggunakan R.

```

gcv2=function(para)
{
  data=read.table("F:/semester AKHIR/TA/wanin/dataTA1.txt",header=T)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  F=matrix(0,nrow=p,ncol=p)
  dataA=data[, (para+2):q]
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot=matrix(ncol=m,nrow=nk)
  for (j in (1:nk))
  {
    a=seq(min(dataA),max(dataA),length.out=50)
    knot[j]=a[j]
  }
  z=(nk*(nk-1)/2)
  knot2=cbind(rep(NA,(z+1)))
  for (i in (1:m))
  {
    knot1=rbind(rep(NA,2))
    for (j in 1:(nk-1))
    {
      for (k in (j+1):nk)
      {
        xx=cbind(knot[j,i],knot[k,i])
        knot1=rbind(knot1,xx)
      }
    }
    knot2=cbind(knot2,knot1)
  }
  knot2=knot2[2:(z+1),2:(2*m+1)]
  aa=rep(1,p)
  data2=matrix(ncol=(2*m),nrow=p)
  data1=data[,2:q]
  a1=length(knot2[,1])
  GCV=rep(NA,a1)
  Rsq=rep(NA,a1)
  for (i in 1:a1)
  {
    for (j in 1:(2*m))
    {
      if(mod(j,2)==1) b=floor(j/2)+1 else b=j/2
      for (k in 1:p)

```



```

{
  if (data1[k,b]<knot2[i,j]) data2[k,j]=0 else data2[k,j]=data1[k,b]-knot2[i,j]
}
}
mx=cbind(aa,data1,data1[,2]^2,data2^2)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in (1:p))
{
  sum=(data[r,1]-yhat[r,])^2
  sum1=(yhat[r,]-mean(data[,1]))^2
  SSE=SSE+sum
  SSR=SSR+sum1
}
Rsq[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsqr=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline kuadratik 2 knot", "\n")
cat("=====", "\n")
cat("=====", "\n")
cat("Rsqr dengan Spline kuadratik 2 knot", "\n")
cat("=====", "\n")
print(Rsqr)
cat("=====", "\n")
cat("HASIL GCV dengan Spline kuadratik 2 knot", "\n")
cat("=====", "\n")
print(GCV)
s1=min(GCV)
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline kuadratik 2 knot", "\n")
cat("=====", "\n")
cat(" GCV =",s1, "\n")
write.table(GCV,file="F:/semester AKHIR/TA/wanin/output GCV2.txt")
write.table(Rsqr,file="F:/semester AKHIR/TA/wanin/output Rsqr2.txt")
write.table(knot2,file="F:/semester AKHIR/TA/wanin/output knot2.txt")
write.table(mx,file="F:/semester AKHIR/TA/wanin/mx.txt")
gcv2(1)

```

Lampiran 8. *Syntax* Pemilihan Titik Knot Optimal dengan Kuadratik Tiga Titik Knot Menggunakan R.

```

gcv3=function(para)
{
  data=read.table("F:/semester AKHIR/TA/wanin/dataTA1.txt",header=T)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  F=matrix(0,nrow=p,ncol=p)
  dataA=data[, (para+2):q]
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot=matrix(ncol=m,nrow=nk)
  for (j in (1:nk))
  {
    a=seq(min(dataA),max(dataA),length.out=50)
    knot[j]=a[j]
  }
  knot=knot[2:(nk-1),]
  knot=data.frame(knot)
  a2=nrow(knot)
  z=(a2*(a2-1)*(a2-2)/6)
  knot1=cbind(rep(NA,(z+1)))
  for (i in (1:m))
  {
    knot2=rbind(rep(NA,3))
    for (j in 1:(a2-2))
    {
      for (k in (j+1):(a2-1))
      {
        for (g in (k+1):a2)
        {
          xx=cbind(knot[j,i],knot[k,i],knot[g,i])
          knot2=rbind(knot2,xx)
        }
      }
    }
    knot1=cbind(knot1,knot2)
  }
  knot1=knot1[2:(z+1),2:(3*m+1)]
  aa=rep(1,p)
  data1=matrix(ncol=(3*m),nrow=p)
  data2=data[, (para+2):q]
  a1=length(knot1[,1])
  GCV=rep(NA,a1)
  Rsq=rep(NA,a1)
}

```

```

data2=data.frame(data2)
for (i in 1:a1)
{
  for (j in 1:ncol(knot1))
  {
    b=ceiling(j/3)
    for (k in 1:p)
    {

      if (data2[k,b]<knot1[i,j]) data1[k,j]=0 else data1[k,j]=data2[k,b]-knot1[i,j]
    }

  }
mx=cbind(aa,data[,2:q],data[,3]^2,data1^2)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in (1:p))
{
  sum=(data[r,1]-yhat[r,])^2
  sum1=(yhat[r,]-mean(data[,1]))^2
  SSE=SSE+sum
  SSR=SSR+sum1
}
Rsq[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsq=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline kuadratik 3 knot", "\n")
cat("=====", "\n")
cat("=====", "\n")
cat("Rsq dengan Spline kuadratik 3 knot", "\n")
cat("=====", "\n")
print (Rsq)
r=max(Rsq)
print (r)
cat("=====", "\n")
cat("HASIL GCV dengan Spline kuadratik 3 knot", "\n")
cat("=====", "\n")
print (GCV)

```

```
s1=min(GCV)
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline kuadrat 3 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.table(GCV, file="F:/semester AKHIR/TA/wanin/output GCV3.txt")
write.table(Rsq, file="F:/semester AKHIR/TA/wanin/output Rsq3.txt")
write.table(knot1, file="F:/semester AKHIR/TA/wanin/output knot3.txt")
}
gcv3(1)
```

Lampiran 9. *Syntax* Pemilihan Titik Knot Optimal dengan Kubik Satu Titik Knot Menggunakan R.

```

library(pracma)
library(MASS)

gcv1=function(para)
{
  data=read.table("F:/semester AKHIR/TA/wanin/dataTA1.txt",header=T)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  dataA=data[, (para+2):q]
  F=matrix(0,nrow=p,ncol=p)
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot1=matrix(ncol=m,nrow=nk)
  for (j in (1:nk))
  {
    a=seq(min(dataA),max(dataA),length.out=50)
    knot1[j]=a[j]
  }
  a1=length(knot1[,1])
  knot1=knot1[2:(a1-1),]
  knot1=data.frame(knot1)
  aa=rep(1,p)
  data1=matrix(ncol=m,nrow=p)
  data2=data[,2:q]
  a2=nrow(knot1)
  GCV=rep(NA,a2)
  Rsq=rep(NA,a2)
  for (i in 1:a2)
  {
    for (j in 1:m)
    {
      for (k in 1:p)
      {
        if(data[k,(j+para+1)]<knot1[i,j]) data1[k,j]=0 else data1[k,j]=data[k,(j+para+1)]-
knot1[i,j]
      }
    }
  }
  mx=cbind(aa,data2,data2[,2]^2,data2[,2]^3,data1^3)
  mx=as.matrix(mx)
  C=pinv(t(mx)%*%mx)
  B=C%*%(t(mx)%*%data[,1])
  yhat=mx%*%B
  SSE=0
}

```

```

SSR=0
for (r in (1:p))
{
  sum=(data[r,1]-yhat[r,])^2
  sum1=(yhat[r,]-mean(data[,1]))^2
  SSE=SSE+sum
  SSR=SSR+sum1
}
Rsq[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsq=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline kubik 1 knot", "\n")
cat("=====", "\n")
cat("=====", "\n")
cat("Rsq dengan Spline kubik 1 knot", "\n")
cat("=====", "\n")
print (Rsq)
cat("=====", "\n")
cat("HASIL GCV dengan Spline kubik 1 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
print(max(Rsq))
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline kubik 1 knot", "\n")
cat("=====", "\n")
cat(" GCV =",s1, "\n")
write.table(GCV,file="F:/semester AKHIR/TA/wanin/output GCV.txt")
write.table(Rsq,file="F:/semester AKHIR/TA/wanin/output Rsq.txt")
write.table(knot1,file="F:/semester AKHIR/TA/wanin/output knot.txt")
}
gcv1(1)

```

Lampiran 10. *Syntax* Pemilihan Titik Knot Optimal dengan Kubik Dua Titik Knot Menggunakan R.

```

gcv2=function(para)
{
  data=read.table("F:/semester AKHIR/TA/wanin/dataTA1.txt",header=T)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  F=matrix(0,nrow=p,ncol=p)
  dataA=data[, (para+2):q]
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot=matrix(ncol=m,nrow=nk)
  for (j in (1:nk))
  {
    a=seq(min(dataA),max(dataA),length.out=50)
    knot[j]=a[j]
  }
  z=(nk*(nk-1)/2)
  knot2=cbind(rep(NA,(z+1)))
  for (i in (1:m))
  {
    knot1=rbind(rep(NA,2))
    for (j in 1:(nk-1))
    {
      for (k in (j+1):nk)
      {
        xx=cbind(knot[j,i],knot[k,i])
        knot1=rbind(knot1,xx)
      }
    }
    knot2=cbind(knot2,knot1)
  }

  knot2=knot2[2:(z+1),2:(2*m+1)]
  aa=rep(1,p)
  data2=matrix(ncol=(2*m),nrow=p)
  data1=data[,2:q]
  a1=length(knot2[,1])
  GCV=rep(NA,a1)
  Rsq=rep(NA,a1)

  for (i in 1:a1)
  {
    for (j in 1:(2*m))
    {

```

```

if(mod(j,2)==1) b=floor(j/2)+1 else b=j/2
for (k in 1:p)
{

  if (data1[k,b]<knot2[i,j]) data2[k,j]=0 else data2[k,j]=data1[k,b]-knot2[i,j]
}

}
mx=cbind(aa,data1,data1[,2]^2,data1[,2]^3,data2^3)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in (1:p))
{
  sum=(data[r,1]-yhat[r,])^2
  sum1=(yhat[r,]-mean(data[,1]))^2
  SSE=SSE+sum
  SSR=SSR+sum1
}
Rsq[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)

A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsq=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline kubik 2 knot", "\n")
cat("=====", "\n")
cat("=====", "\n")
cat("Rsq dengan Spline kubik 2 knot", "\n")
cat("=====", "\n")
print(Rsq)
cat("=====", "\n")
cat("HASIL GCV dengan Spline kubik 2 knot", "\n")
cat("=====", "\n")
print(GCV)
s1=min(GCV)
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline kubik 2 knot", "\n")
cat("=====", "\n")
cat(" GCV =",s1, "\n")
write.table(GCV,file="F:/semester AKHIR/TA/wanin/output GCV2.txt")

```



```
write.table(Rsq,file="F:/semester AKHIR/TA/wanin/output Rsq2.txt")
write.table(knot2,file="F:/semester AKHIR/TA/wanin/output knot2.txt")
write.table(mx,file="F:/semester AKHIR/TA/wanin/mx.txt")
}
gcv2(1)
```

Lampiran 11. *Syntax* Pemilihan Titik Knot Optimal dengan Kubik Tiga Titik Knot Menggunakan R.

```

gev3=function(para)
{
  data=read.table("F:/semester AKHIR/TA/wanin/dataTA1.txt",header=T)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  F=matrix(0,nrow=p,ncol=p)
  dataA=data[, (para+2):q]
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot=matrix(ncol=m,nrow=nk)
  for (j in (1:nk))
  {
    a=seq(min(dataA),max(dataA),length.out=50)
    knot[j]=a[j]
  }
  knot=knot[2:(nk-1),]
  knot=data.frame(knot)
  a2=nrow(knot)
  z=(a2*(a2-1)*(a2-2))/6
  knot1=cbind(rep(NA,(z+1)))
  for (i in (1:m))
  {
    knot2=rbind(rep(NA,3))
    for (j in 1:(a2-2))
    {
      for (k in (j+1):(a2-1))
      {
        for (g in (k+1):a2)
        {
          xx=cbind(knot[j,i],knot[k,i],knot[g,i])
          knot2=rbind(knot2,xx)
        }
      }
    }
    knot1=cbind(knot1,knot2)
  }
  knot1=knot1[2:(z+1),2:(3*m+1)]
  aa=rep(1,p)
  data1=matrix(ncol=(3*m),nrow=p)
  data2=data[, (para+2):q]
  a1=length(knot1[,1])
  GCV=rep(NA,a1)
  Rsq=rep(NA,a1)
}

```

```

data2=data.frame(data2)
for (i in 1:a1)
{
  for (j in 1:ncol(knot1))
  {
    b=ceiling(j/3)
    for (k in 1:p)
    {
      if (data2[k,b]<knot1[i,j]) data1[k,j]=0 else data1[k,j]=data2[k,b]-knot1[i,j]
    }
  }
mx=cbind(aa,data[,2:q],data[,3]^2,data[,3]^3,data1^3)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in (1:p))
{
  sum=(data[r,1]-yhat[r,])^2
  sum1=(yhat[r,]-mean(data[,1]))^2
  SSE=SSE+sum
  SSR=SSR+sum1
}
Rsq[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsq=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline kubik 3 knot", "\n")
cat("=====", "\n")
cat("=====", "\n")
cat("Rsq dengan Spline kubik 3 knot", "\n")
cat("=====", "\n")
print (Rsq)
r=max(Rsq)
print (r)
cat("=====", "\n")
cat("HASIL GCV dengan Spline kubik 3 knot", "\n")
cat("=====", "\n")
print (GCV)

```

```
s1=min(GCV)
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline kubik 3 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.table(GCV, file="F:/semester AKHIR/TA/wanin/output GCV3.txt")
write.table(Rsq, file="F:/semester AKHIR/TA/wanin/output Rsq3.txt")
write.table(knot1, file="F:/semester AKHIR/TA/wanin/output knot3.txt")

}
gcv3(1)
```

Lampiran 12. Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Linier Satu Titik Knot

No	Knot 1	GCV
1	4.037347	1.294646
2	4.524694	1.291593
3	5.012041	1.293446
4	5.499388	1.295013
5	5.986735	1.295014
6	6.474082	1.292091
7	6.961429	1.291426
8	7.448776	1.287464
9	7.936122	1.285698
10	8.423469	1.284667
11	8.910816	1.284192
12	9.398163	1.287467
13	9.88551	1.290145
14	10.37286	1.292123
15	10.8602	1.293511
16	11.34755	1.294317
17	11.8349	1.295298
18	12.32224	1.29545
19	12.80959	1.295539
20	13.29694	1.295567
...
44	24.99327	1.252053
45	25.48061	1.252053
46	25.96796	1.252053
47	26.45531	1.252053
48	26.94265	1.252053

Lampiran 13. Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Linier Dua Titik Knot

No	Knot 1	Knot 2	GCV
1	3.55	4.037347	1.21333
2	3.55	4.524694	1.21333
3	3.55	5.012041	1.21333
4	3.55	5.499388	1.21333
5	3.55	5.986735	1.21333
6	3.55	6.474082	1.21333
7	3.55	6.961429	1.21333
8	3.55	7.448776	1.21333
9	3.55	7.936122	1.295508
10	3.55	8.423469	1.295508
11	3.55	8.910816	1.295508
12	3.55	9.398163	1.295508
13	3.55	9.88551	1.295508
14	3.55	10.37286	1.295508
15	3.55	10.8602	1.295508
16	3.55	11.34755	1.295508
17	3.55	11.8349	1.295508
18	3.55	12.32224	1.295508
19	3.55	12.80959	1.295508
20	3.55	13.29694	1.295508
...
1221	25.96796	26.94265	1.374349
1222	25.96796	27.43	1.374349
1223	26.45531	26.94265	1.374349
1224	26.45531	27.43	1.374349
1225	26.94265	27.43	1.374349

Lampiran 14. Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Linier Tiga Titik Knot

No	Knot 1	Knot 2	Knot 3	GCV
1	4.037347	4.524694	5.012041	1.379447
2	4.037347	4.524694	5.499388	1.384729
3	4.037347	4.524694	5.986735	1.393654
4	4.037347	4.524694	6.474082	1.402957
5	4.037347	4.524694	6.961429	1.403383
6	4.037347	4.524694	7.448776	1.402018
7	4.037347	4.524694	7.936122	1.400591
8	4.037347	4.524694	8.423469	1.399546
9	4.037347	4.524694	8.910816	1.39886
10	4.037347	4.524694	9.398163	1.40105
11	4.037347	4.524694	9.88551	1.402436
12	4.037347	4.524694	10.37286	1.403144
13	4.037347	4.524694	10.8602	1.403396
14	4.037347	4.524694	11.34755	1.403381
15	4.037347	4.524694	11.8349	1.40276
16	4.037347	4.524694	12.32224	1.402491
17	4.037347	4.524694	12.80959	1.402203
18	4.037347	4.524694	13.29694	1.401872
19	4.037347	4.524694	13.78429	1.402033
20	4.037347	4.524694	14.27163	1.40267
...
17292	24.99327	26.45531	26.94265	1.252053
17293	25.48061	25.96796	26.45531	1.252053
17294	25.48061	25.96796	26.94265	1.252053
17295	25.48061	26.45531	26.94265	1.252053
17296	25.96796	26.45531	26.94265	1.252053

Lampiran 15. Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Kuadratik Satu Titik Knot

No	Knot 1	GCV
1	4.037347	1.365532
2	4.524694	1.377774
3	5.012041	1.378161
4	5.499388	1.378243
5	5.986735	1.37809
6	6.474082	1.378074
7	6.961429	1.378193
8	7.448776	1.378243
9	7.936122	1.378151
10	8.423469	1.37795
11	8.910816	1.377683
12	9.398163	1.377522
13	9.88551	1.37761
14	10.37286	1.377855
15	10.8602	1.378119
16	11.34755	1.378243
17	11.8349	1.378056
18	12.32224	1.37735
19	12.80959	1.376033
20	13.29694	1.373952
...
44	24.99327	1.328745
45	25.48061	1.328745
46	25.96796	1.328745
47	26.45531	1.328745
48	26.94265	1.328745

Lampiran 16. Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Kuadratik Dua Titik Knot

No	Knot 1	Knot 2	GCV
1	3.55	4.037347	1.269648
2	3.55	4.524694	1.269648
3	3.55	5.012041	1.269648
4	3.55	5.499388	1.269648
5	3.55	5.986735	1.269648
6	3.55	6.474082	1.269648
7	3.55	6.961429	1.269648
8	3.55	7.448776	1.269648
9	3.55	7.936122	1.269646
10	3.55	8.423469	1.267835
11	3.55	8.910816	1.263565
12	3.55	9.398163	1.257864
13	3.55	9.88551	1.251434
14	3.55	10.37286	1.244762
15	3.55	10.8602	1.238183
16	3.55	11.34755	1.231925
17	3.55	11.8349	1.226145
18	3.55	12.32224	1.220943
19	3.55	12.80959	1.21638
20	3.55	13.29694	1.212488
...
1221	25.96796	26.94265	1.263067
1222	25.96796	27.43	1.2617
1223	26.45531	26.94265	1.261721
1224	26.45531	27.43	1.260429
1225	26.94265	27.43	1.259221

Lampiran 17. Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Kuadratik Tiga Titik Knot

No	Knot 1	Knot 2	Knot 3	GCV
1	4.037347	4.524694	5.012041	1.46641
2	4.037347	4.524694	5.499388	1.480699
3	4.037347	4.524694	5.986735	1.478209
4	4.037347	4.524694	6.474082	1.478776
5	4.037347	4.524694	6.961429	1.479995
6	4.037347	4.524694	7.448776	1.480607
7	4.037347	4.524694	7.936122	1.480824
8	4.037347	4.524694	8.423469	1.480626
9	4.037347	4.524694	8.910816	1.480261
10	4.037347	4.524694	9.398163	1.480047
11	4.037347	4.524694	9.88551	1.480214
12	4.037347	4.524694	10.37286	1.48056
13	4.037347	4.524694	10.8602	1.480814
14	4.037347	4.524694	11.34755	1.480713
15	4.037347	4.524694	11.8349	1.479962
16	4.037347	4.524694	12.32224	1.478292
17	4.037347	4.524694	12.80959	1.475675
18	4.037347	4.524694	13.29694	1.47194
19	4.037347	4.524694	13.78429	1.467276
20	4.037347	4.524694	14.27163	1.46211
...
17292	24.99327	26.45531	26.94265	1.328745
17293	25.48061	25.96796	26.45531	1.328745
17294	25.48061	25.96796	26.94265	1.328745
17295	25.48061	26.45531	26.94265	1.328745
17296	25.96796	26.45531	26.94265	1.328745

Lampiran 18. Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Kubik Satu Titik Knot

No	Knot 1	GCV
1	4.037347	1.364078
2	4.524694	1.363943
3	5.012041	1.36223
4	5.499388	1.359715
5	5.986735	1.357458
6	6.474082	1.355583
7	6.961429	1.353383
8	7.448776	1.35092
9	7.936122	1.347835
10	8.423469	1.344145
11	8.910816	1.436091
12	9.398163	1.430104
13	9.88551	1.423788
14	10.37286	1.417464
15	10.8602	1.411281
16	11.34755	1.405288
17	11.8349	1.399607
18	12.32224	1.394632
19	12.80959	1.390477
20	13.29694	1.387207
...
44	24.99327	1.406425
45	25.48061	1.406425
46	25.96796	1.406425
47	26.45531	1.406425
48	26.94265	1.363189

Lampiran 19. Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Kubik Dua Titik Knot

No	Knot 1	Knot 2	GCV
1	3.55	4.037347	1.34626
2	3.55	4.524694	1.348527
3	3.55	5.012041	1.350946
4	3.55	5.499388	1.353525
5	3.55	5.986735	1.356272
6	3.55	6.474082	1.35919
7	3.55	6.961429	1.362286
8	3.55	7.448776	1.365564
9	3.55	7.936122	1.369031
10	3.55	8.423469	1.372683
11	3.55	8.910816	1.376482
12	3.55	9.398163	1.380397
13	3.55	9.88551	1.384408
14	3.55	10.37286	1.3885
15	3.55	10.8602	1.392664
16	3.55	11.34755	1.396893
17	3.55	11.8349	1.401182
18	3.55	12.32224	1.405529
19	3.55	12.80959	1.409932
20	3.55	13.29694	1.414391
...
1221	25.96796	26.94265	1.872264
1222	25.96796	27.43	1.881581
1223	26.45531	26.94265	1.881445
1224	26.45531	27.43	1.890856
1225	26.94265	27.43	1.900255

Lampiran 20. Nilai GCV Beserta Nilai Titik-Titik Knot dengan Menggunakan Kubik Tiga Titik Knot

No	Knot 1	Knot 2	Knot 3	GCV
1	4.037347	4.524694	5.012041	1.456446
2	4.037347	4.524694	5.499388	1.456361
3	4.037347	4.524694	5.986735	1.457722
4	4.037347	4.524694	6.474082	1.458316
5	4.037347	4.524694	6.961429	1.457403
6	4.037347	4.524694	7.448776	1.455622
7	4.037347	4.524694	7.936122	1.452769
8	4.037347	4.524694	8.423469	1.44897
9	4.037347	4.524694	8.910816	1.444322
10	4.037347	4.524694	9.398163	1.438873
11	4.037347	4.524694	9.88551	1.433058
12	4.037347	4.524694	10.37286	1.4272
13	4.037347	4.524694	10.8602	1.421455
14	4.037347	4.524694	11.34755	1.415879
15	4.037347	4.524694	11.8349	1.410596
16	4.037347	4.524694	12.32224	1.405992
17	4.037347	4.524694	12.80959	1.402181
18	4.037347	4.524694	13.29694	1.399225
19	4.037347	4.524694	13.78429	1.397185
20	4.037347	4.524694	14.27163	1.395989
...
17292	24.99327	26.45531	26.94265	1.406425
17293	25.48061	25.96796	26.45531	1.406425
17294	25.48061	25.96796	26.94265	1.406425
17295	25.48061	26.45531	26.94265	1.406425
17296	25.96796	26.45531	26.94265	1.406425

Lampiran 21. *Syntax* Estimasi dan Pengujian Parameter Menggunakan R

```

uji=function(alpha,para)
{
  data=read.table("F:/semester AKHIR/TA/wanin/dataTA1.txt",header=T)
  knot=read.table("F:/semester AKHIR/TA/wanin/UJI PARAMETER.txt")
  data=as.matrix(data)
  knot=as.matrix(knot)
  ybar=mean(data[,1])
  m=para+2
  p=nrow(data)
  q=ncol(data)
  dataA=cbind(data[,m],data[,m],data[,m])
  dataA=as.matrix(dataA)
  satu=rep(1,p)
  n1=ncol(knot)
  data.knot=matrix(ncol=n1,nrow=p)
  for (i in 1:n1)
  {
    for(j in 1:p)
    {
      if(dataA[j,i]<knot[1,i]) data.knot[j,i]=0 else data.knot[j,i]=dataA[j,i]-knot[1,i]
    }
  }
  print(data.knot)
  mx=cbind(satu,data[,2],
           data[,3],data.knot[,1:3])
  mx=as.matrix(mx)
  B=(pinv(t(mx)%*%mx))%*%t(mx)%*%data[,1]
  cat("=====", "\n")
  cat("Estimasi Parameter", "\n")
  cat("=====", "\n")
  print (B)
  n1=nrow(B)
  yhat=mx%*%B
  res=data[,1]-yhat
  SSE=sum((data[,1]-yhat)^2)
  SSR=sum((yhat-ybar)^2)
  SST=SSR+SSE
  MSE=SSE/(p-n1)
  MSR=SSR/(n1-1)
  Rsq=(SSR/(SSR+SSE))*100
  #uji F (uji serentak)
  Fhit=MSR/MSE
  pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
  if (pvalue<=alpha)
  {

```

```

cat("-----", "\n")
cat("Kesimpulan hasil uji serentak", "\n")
cat("-----", "\n")
cat("Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan", "\n")
cat("", "\n")
}
else
{
cat("-----", "\n")
cat("Kesimpulan hasil uji serentak", "\n")
cat("-----", "\n")
cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh signifikan", "\n")
cat("", "\n")
}
#uji t (uji individu)
thit=rep(NA,n1)
pval=rep(NA,n1)
SE=sqrt(diag(MSE*(pinv(t(mx)%*%mx)))
cat("-----", "\n")
cat("Kesimpulan hasil uji individu", "\n")
cat("-----", "\n")
thit=rep(NA,n1)
pval=rep(NA,n1)
for (i in 1:n1)
{
thit[i]=B[i,1]/SE[i]
pval[i]=2*(pt(abs(thit[i]),(p-n1),lower.tail=FALSE))
if (pval[i]<=alpha) cat("Tolak Ho yakni prediktor signifikan dengan
pvalue", pval[i], "\n") else cat("Gagal tolak Ho yakni prediktor tidak signifikan dengan
pvalue", pval[i], "\n")
}
}
thit=as.matrix(thit)
cat("===== ", "\n")
cat("nilai t hitung", "\n")
cat("===== ", "\n")
print (thit)
cat("-----", "\n")
cat("Analysis of Variance", "\n")
cat("===== ", "\n")
cat("Sumber      df      SS      MS      Fhit", "\n")
cat("Regresi      ",(n1-1), " ",SSR," ",MSR," ",Fhit," \n")
cat("Error         ",p-n1, " ",SSE," ",MSE," \n")
cat("Total         ",p-1, " ",SST," \n")
cat("===== ", "\n")
cat("s=",sqrt(MSE)," Rsq=",Rsq," \n")
cat("pvalue(F)=", pvalue," \n")
write.table(res,file="F:/semester AKHIR/TA/wanin/output uji residual.txt")
write.table(pval,file="F:/semester AKHIR/TA/wanin/output uji pvalue.txt")
write.table(mx,file="F:/semester AKHIR/TA/wanin/output uji mx.txt")

```

```
write.table(yhat,file="F:/semester AKHIR/TA/wanin/output uji yhat.txt")
write.table(mx,file="F:/semester AKHIR/TA/wanin/mx2.txt")
}
uji(0.05,1)
```


Lampiran 22. Output Estimasi dan Pengujian Parameter Menggunakan R

```

=====
Estimasi Parameter
=====
          [,1]
[1,]  4.93807718
[2,] -0.03547447
[3,]  0.13788756
[4,] -8.84789060
[5,] 20.28787622
[6,] -11.60790849
-----
Kesimpulan hasil uji serentak
-----
Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan
-----
Kesimpulan hasil uji individu
-----
Tolak Ho yakni prediktor signifikan dengan pvalue 1.287118e-
07
Tolak Ho yakni prediktor signifikan dengan pvalue 0.00042982
9
Tolak Ho yakni prediktor signifikan dengan pvalue 0.02502268
Tolak Ho yakni prediktor signifikan dengan pvalue 0.00912449
4
Tolak Ho yakni prediktor signifikan dengan pvalue 0.00468358
1
Tolak Ho yakni prediktor signifikan dengan pvalue 0.00383527
8
=====
nilai t hitung
=====
          [,1]
[1,]  7.003927
[2,] -3.991554
[3,]  2.368047
[4,] -2.801392
[5,]  3.073165
[6,] -3.152874
-----
Analysis of Variance
=====
Sumber   df      SS      MS      Fhit
Regresi   5  51.23765  10.24753  11.33903
Error    28  25.30471   0.9037395
Total    33  76.54235
=====
s= 0.9506521      Rsq= 66.94026
pvalue(F)= 4.798364e-06

```

Lampiran 23. *Output Residual Menggunakan R*

No	residual	No	residual
1	-0.1799	18	0.818389
2	1.047629	19	0.889666
3	0.492815	20	0.328018
4	0.290115	21	0.54832
5	-0.42158	22	0.994558
6	0.548645	23	-1.05679
7	0.029832	24	-1.55276
8	-1.14133	25	1.732295
9	0.094786	26	-0.75045
10	-0.67849	27	0.629409
11	0.274069	28	0.706255
12	-1.31288	29	0.167145
13	-0.82794	30	1.382796
14	-1.40985	31	-0.62628
15	-0.36786	32	1.357165
16	-0.42307	33	0.025637
17	-1.30228	34	-0.3061

Lampiran 24. Syntax Pengujian Glejser Menggunakan R

```

glejser=function(data,knot,res,alpha,para)
{
  data=read.table("F:/semester AKHIR/TA/wanin/dataTA1.txt", sep='\t', header=TRUE)
  knot=read.table("F:/semester AKHIR/TA/wanin/UJI PARAMETER.txt", sep='\t')
  res=read.table("F:/semester AKHIR/TA/wanin/output uji residual.txt")
  data=as.matrix(data)
  knot=as.matrix(knot)
  res=abs(res)
  res=as.matrix(res)
  rbar=mean(res)
  m=para+2
  p=nrow(data)
  q=ncol(data)
  dataA=cbind(data[,m],data[,m],data[,m]) #m
  dataA=as.matrix(dataA)
  satu=rep(1,p)
  n1=ncol(knot)
  data.knot=matrix(ncol=n1,nrow=p)
  for (i in 1:n1)
  {
    for(j in 1:p)
    {
      if (dataA[j,i]<knot[1,i]) data.knot[j,i]=0 else data.knot[j,i]=dataA[j,i]-knot[1,i]
    }
  }
  mx=cbind(satu,data[,2],
           data[,3],data.knot[,1:3])
  mx=as.matrix(mx)
  B=(ginv(t(mx)%*%mx))%*%t(mx)%*%res
  n1=nrow(B)
  yhat=mx%*%B
  residual=res-yhat
  SSE=sum((res-yhat)^2)
  SSR=sum((yhat-rbar)^2)
  SST=SSR+SSE
  MSE=SSE/(p-n1)
  MSR=SSR/(n1-1)
  Rsq=(SSR/SST)*100

  #uji F (uji serentak)
  Fhit=MSR/MSE
  pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
  if (pvalue<=alpha)
  {
    cat("-----","\n")
    cat("Kesimpulan hasil uji serentak","\n")
  }
}

```

```

    cat("-----","\n")
    cat("Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan atau terjadi
heteroskedastisitas","\n")
    cat("","\n")
}
else
{
    cat("-----","\n")
    cat("Kesimpulan hasil uji serentak","\n")
    cat("-----","\n")
    cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh signifikan atau tidak
terjadi heteroskedastisitas","\n")
    cat("","\n")
}
cat("Analysis of Variance","\n")
cat("=====","\n")
cat("Sumber    df    SS    MS    Fhit","\n")
cat("Regresi    ,(n1-1), " ",SSR," ",MSR," ",Fhit,"\n")
cat("Error      ,p-1," " ",SSE," ",MSE,"\n")
cat("Total      ,p-1," " ",SST,"\n")
cat("=====","\n")
cat("s=",sqrt(MSE)," Rsq=",Rsq,"\n")
cat("pvalue(F)=",pvalue,"\n")
}
glejser(data,knot,res,0.05,1)

```

Lampiran 25. Output Uji Glejser Menggunakan R

 Kesimpulan hasil uji serentak

Gagal Tolak H_0 yakni semua prediktor tidak berpengaruh signifikan atau tidak terjadi heteroskedastisitas

Analysis of Variance

Sumber	df	SS	MS	Fhit
Regresi	5	1.691236	0.3382472	1.676948
Error	28	5.647713	0.201704	
Total	33	7.33895		

=====

s= 0.4491147 Rsq= 23.04466
 pvalue(F)= 0.1728965

Lampiran 26. Surat Keterangan Pengambilan Data**SURAT PERNYATAAN**

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FSAD ITS:

Nama : Wanindyatami Firstidi Putri

NRP : 062115 4000 0037

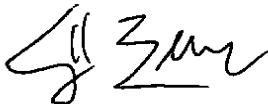
menyatakan bahwa data yang digunakan dalam Tugas Akhir/ Thesis ini merupakan data sekunder yang diambil dari penelitian / buku/ Tugas Akhir/ Thesis/ publikasi lainnya yaitu:

Sumber : Profil Kesehatan Indonesia tahun 2018

Keterangan : Data persentase balita gizi buruk di Indonesia

Surat Pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui
Pembimbing Tugas Akhir



Jerry Dwi Trijoyo P, S.Si., M.Si., Ph.D
NIP. 19810223 200812 1 003

Surabaya, 7 Januari 2020



Wanindyatami Firstidi P
NRP.062115 4000 0037

BIODATA PENULIS



Penulis dengan nama lengkap Wanidyatami Firstidi Putri dan sering dipanggil Nindi atau Wanin dilahirkan di Pati pada tanggal 13 Januari 1997 dari pasangan Bapak Budi Santosa dan Ibu Triyanti. Penulis menempuh Pendidikan formal di SDN Kauman 01, SMPN 1 Juwana, dan SMAN 1 Pati. Setelah lulus SMA penulis diterima sebagai Mahasiswa Departemen Statistika ITS melalui jalur undangan SNMPTN pada tahun 2015. Anak ke-satu dari dua bersaudara. Penulis mempunyai adik yang bernama Satrya Pinandhita Scondhitya Putra. Pada bulan Juni-Juli 2018 penulis berkesempatan untuk melakukan *internship program* di PT. Dua Kelinci Kab. Pati. Apabila pembaca ingin memberi kritik dan saran serta ingin berdiskusi lebih lanjut mengenai Tugas Akhir ini, dapat menghubungi penulis melalui email wanindyatami@gmail.com, nomor telepon 082139436197, dan no WA 085876879267.

(Halaman ini sengaja dikosongkan)