



TUGAS AKHIR – SS141501

**PEMODELAN FAKTOR-FAKTOR YANG
MEMENGARUHI INDEKS PEMBANGUNAN
MANUSIA DI PROVINSI SUMATERA UTARA
MENGUNAKAN REGRESI NONPARAMETRIK
SPLINE**

**NICEA ROONA PARANOAN
NRP 1313 100 125**

**Dosen Pembimbing
Prof. Dr. Drs I Nyoman Budiantara, M. Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**



TUGAS AKHIR – SS 141501

**PEMODELAN FAKTOR-FAKTOR YANG
MEMENGARUHI INDEKS PEMBANGUNAN
MANUSIA DI PROVINSI SUMATERA UTARA
MENGUNAKAN REGRESI NONPARAMETRIK
SPLINE**

**NICEA ROONA PARANOAN
NRP 1313 100 125**

**Dosen Pembimbing
Prof. Dr. Drs. I Nyoman Budiantara, M. Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**



FINAL PROJECT – SS 141501

***MODELING FACTORS THAT INFLUENCE
HUMAN DEVELOPMENT INDEX
IN NORTH SUMATERA USING
NONPARAMETRIC SPLINE REGRESSION***

**NICEA ROONA PARANOAN
NRP 1313 100 125**

Supervisor

Prof. Dr. Drs. I Nyoman Budiantara, M. Si

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**

LEMBAR PENGESAHAN

PEMODELAN FAKTOR-FAKTOR YANG MEMENGARUHI INDEKS PEMBANGUNAN MANUSIA DI PROVINSI SUMATERA UTARA MENGUNAKAN REGRESI NONPARAMETRIK SPLINE

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada

Program Studi Sarjana Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh :

Nicea Roona Paranoan

NRP. 1313 100 125

Disetujui oleh Pembimbing:

Prof. Dr. Drs. I Nyoman Budiantara, M.Si

NIP. 19650603 198903 1 003

Mengetahui,
Kepala Departemen




Dr. Suhartono
NIP. 19710929 199512 1 001

SURABAYA, JULI 2017

**PEMODELAN FAKTOR-FAKTOR YANG
MEMENGARUHI INDEKS PEMBANGUNAN MANUSIA
DI PROVINSI SUMATERA UTARA MENGGUNAKAN
REGRESI NONPARAMETRIK SPLINE**

Nama Mahasiswa : Nicea Roona Paranoan
NRP : 1313 100 125
Departemen : Statistika
**Dosen Pembimbing : Prof. Dr. Drs. I Nyoman
Budiantara, M.Si**

Abstrak

Indeks Pembangunan Manusia (IPM) merupakan indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia dan mampu menjelaskan bagaimana manusia dapat mengakses hasil pembangunan dalam dimensi kesehatan, pendidikan serta standar hidup layak. Sumatera Utara adalah salah satu provinsi di Indonesia yang mengalami kemajuan dalam pembangunan manusia setiap tahunnya namun angka IPM Sumatera Utara masih di bawah angka nasional yang mencapai 69,55 dan terdapat beberapa wilayah yang masih memiliki angka IPM rendah. Terdapat 6 faktor yang memengaruhi IPM, yaitu Angka Partisipasi Murni SD, Angka Partisipasi Murni SMP, Angka Partisipasi Murni SMA, Tingkat Partisipasi Angkatan Kerja, Tingkat Pengangguran Terbuka dan Angka Harapan Hidup. Pemodelan IPM menggunakan regresi nonparametrik Spline karena hubungan antara IPM dengan faktor-faktor yang memengaruhinya tidak membentuk pola tertentu. Berdasarkan hasil analisis, model terbaik menggunakan kombinasi titik knot (2,1,2,3,3,3) dan seluruh faktor berpengaruh signifikan terhadap IPM dengan koefisien determinasi sebesar 91,80 persen.

Kata kunci : Indeks Pembangunan Manusia, regresi nonparametrik spline, titik knot.

(Halaman ini sengaja dikosongkan)

MODELING FACTORS THAT INFLUENCE HUMAN DEVELOPMENT INDEX IN NORTH SUMATERA USING NONPARAMETRIC SPLINE REGRESSION

Student's Name : Nicea Roona Paranoan
NRP : 1313 100 125
Department : Statistics
Supervisor : Prof. Dr. Drs. I Nyoman Budiantara, M.Si

Abstract

Human Development Index (HDI) is an important indicator to measure the success an effort to build the quality of human life and explain how people can access the results of development in the dimensions of health, education and decent living standards. North Sumatra is one of Provinces in Indonesia that progress in human development. Although North Sumatra have increased HDI, North Sumatra's HDI is still below the national average of 69.55 and there are some areas that still have a low HDI. There are 6 factors suspected to affect the HDI, namely Net Enrolment Rate of Elementary, Net Enrolment Rate of Junior High, Net Enrolment Rate of Senior High, Labour Force Participation Rate, Open Employee Rate and Life Expectancy. Modeling IPM using nonparametric Spline regression because HDI pattern and factors that influence HDI in North Sumatra does not have a certain pattern. Based on the analysis, the best nonparametric Spline regression model using knots combination of (2,1,2,3,3,3) and all factors significant effect on the HDI then the model have R^2 of 91.80 percent.

Keywords : *Human Development Index, knots, nonparametric regression, spline.*

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur kehadiran Tuhan Yang Maha Esa, atas segala limpah rahmat dan karunia-Nya yang tak pernah berhenti diberikan, sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul **“Pemodelan Faktor-Faktor yang Memengaruhi Indeks Pembangunan Manusia di Provinsi Sumatera Utara Menggunakan Regresi Nonparametrik Spline”** dengan baik.

Penulis menyadari bahwa dalam penyusunan Tugas Akhir ini tidak lepas dari banyaknya bantuan dan dukungan yang diberikan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis mengucapkan terima kasih kepada:

1. Prof. Dr. Drs. I Nyoman Budiantara, M.Si selaku dosen pembimbing yang telah sabar dalam memberikan ilmu, bimbingan, saran, dan dukungan selama penyusunan Tugas Akhir.
2. Ibu Dr. Vita Ratnasari, S.Si, M.Si dan Ibu Dr. Ismaini Zain, M.Si selaku dosen penguji yang telah memberikan banyak bantuan dan masukan untuk kesempurnaan Tugas Akhir ini.
3. Bapak Pasang Paranoan, Ibu Yospina Kamma, Kakak Laura Evalina Paranoan dan Adik Arthur Anugerah Paranoan sebagai orang tua dan saudara penulis yang selalu memberikan doa, semangat, serta motivasi bagi penulis dalam menyelesaikan Tugas Akhir.
4. Dr. Suhartono selaku Kepala Departemen Statistika ITS dan Dr. Sutikno, S.Si, M.Si selaku Ketua Program Studi Sarjana Statistika ITS.

Penulis berharap hasil Tugas Akhir ini dapat bermanfaat bagi semua pihak. Penulis menyadari masih ada kekurangan dalam penulisan Tugas Akhir ini, untuk itu kritik dan saran dari semua pihak sangat diharapkan.

Surabaya, 13 Juli 2017

Penulis

(Halaman ini sengaja dikosongkan)

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR TABEL	xv
DAFTAR GAMBAR	xvii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian	5
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah	5
BAB II TINJAUAN PUSTAKA	
2.1 Statistika Deskriptif.....	7
2.2 Analisis Regresi	7
2.3 Regresi Nonparametrik Spline	8
2.4 Pemilihan Titik Knot Optimal.....	9
2.5 Kriteria Pemilihan Pemodelan Terbaik.....	10
2.6 Pengujian Parameter Model	10
2.6.1 Pengujian Secara Serentak.....	10
2.6.2 Pengujian Secara Parsial.....	11
2.7 Pengujian Asumsi Residual	12
2.7.1 Uji Asumsi Identik.....	12
2.7.2 Uji Asumsi Independen	13
2.7.3 Uji Asumsi Distribusi Normal	15
2.8 Indeks Pembangunan Manusia	16
2.9 Kerangka Konsep	18

BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data.....	21
3.2 Variabel Penelitian	21
3.3 Struktur Data	24
3.4 Langkah Analisis.....	24
BAB IV ANALISIS DAN PEMBAHASAN	
4.1 Karakteristik Data Indeks Pembangunan Manusia	27
4.2 Pemodelan Indeks Pembangunan Manusia.....	34
4.2.1 Analisis Faktor-Faktor yang Memengaruhi Indeks Pembangunan Manusia.....	35
4.2.2 Pemilihan Titik Knot Optimum.....	36
4.2.3 Pengujian Signifikansi Parameter Model Regresi Nonparametrik Spline.....	43
4.2.3 Pengujian Asumsi Residual.....	46
4.3 Pembahasan.....	49
BAB V KESIMPULAN DAN SARAN	
5.1 Kesimpulan	55
5.2 Saran.....	56
DAFTAR PUSTAKA	57
LAMPIRAN	61
BIODATA PENULIS	95

DAFTAR TABEL

Tabel 2.1	Penentuan Status IPM Berdasarkan Nilai IPM.....	17
Tabel 3.1	Variabel Penelitian	21
Tabel 3.2	Struktur Data Penelitian.....	24
Tabel 4.1	Jumlah Kabupaten/Kota Berdasarkan Status IPM ...	27
Tabel 4.2	Kabupaten/Kota yang Terkategorikan dalam Status IPM Sedang	28
Tabel 4.3	Kabupaten/Kota yang Terkategorikan dalam Status IPM Tinggi.....	29
Tabel 4.4	Statistika Deskriptif dari Variabel Penelitian	29
Tabel 4.5	GCV 1 Titik Knot	37
Tabel 4.6	GCV 2 Titik Knot.....	38
Tabel 4.7	GCV 3 Titik Knot.....	40
Tabel 4.8	GCV Kombinasi Titik Knot	42
Tabel 4.9	Perbandingan GCV Berbagai Knot	43
Tabel 4.10	Tabel ANOVA Pengujian Secara Serentak	44
Tabel 4.11	Hasil Pengujian Secara Parsial	45
Tabel 4.12	ANOVA Uji Glejser	46
Tabel 4.13	Daftar Kabupaten/Kota dalam Interval Pertama x_5 ..	53

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

Gambar 3.1 Diagram Alir	26
Gambar 4.1 <i>Scatterplot</i> antara IPM dengan 6 Variabel Prediktor.....	35
Gambar 4.2 Plot ACF Residual	47
Gambar 4.3 <i>Normal Probability Plot</i> Residual.....	48

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

Halaman

Lampiran 1	Data Indeks Pembangunan Manusia Sumatera Utara dan Faktor-Faktor yang Diduga Memengaruhinya.....	61
Lampiran 2	Program Pemilihan Titik Knot Optimal dengan Satu Titik Knot Menggunakan <i>Software R</i>	63
Lampiran 3	Program Pemilihan Titik Knot Optimal dengan Dua Titik Knot Menggunakan <i>Software R</i>	66
Lampiran 4	Program Pemilihan Titik Knot Optimal dengan Tiga Titik Knot Menggunakan <i>Software R</i>	69
Lampiran 5	Program Pemilihan Titik Knot Optimal dengan Kombinasi Titik Knot Menggunakan <i>Software R</i>	72
Lampiran 6	Program Estimasi Parameter dengan Kombinasi Titik Knot 2,1,2,3,3,3	80
Lampiran 7	Program Uji Glejser untuk Kombinasi Titik Knot 2,1,2,3,3,3	83
Lampiran 8	<i>Output</i> Nilai GCV dengan Satu Titik Knot	85
Lampiran 9	<i>Output</i> Nilai GCV dengan Dua Titik Knot	86
Lampiran 10	<i>Output</i> Nilai GCV dengan Tiga Titik Knot	87
Lampiran 11	<i>Output</i> Estimasi Parameter dan Uji Signifikansi.....	89
Lampiran 12	<i>Output</i> Uji Glejser	92
Lampiran 13	Surat Pernyataan Data	93

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Peningkatan kesejahteraan masyarakat di suatu negara atau wilayah membutuhkan pembangunan manusia. *United Nation Development Programme* (UNDP) mendefinisikan pembangunan subpenduduk. Berdasarkan konsep tersebut penduduk ditempatkan sebagai tujuan akhir (*the ultimated end*), sedangkan upaya pembangunan dipandang sebagai sarana (*principal means*) untuk mencapai tujuan itu. Ada beberapa hal pokok yang perlu diperhatikan untuk menjamin tercapainya tujuan pembangunan manusia adalah produktivitas, pemerataan, kesinambungan, pemberdayaan (BPS, 2016). Berbagai macam ukuran pembangunan manusia dibuat, namun tidak semua digunakan sebagai ukuran standar yang dapat membandingkan pembangunan manusia antar negara atau wilayah. Oleh karena itu, Badan Perserikatan Bangsa-Bangsa (PBB) menetapkan suatu ukuran standar yaitu Indeks Pembangunan Manusia (IPM). IPM merupakan indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia. IPM menjelaskan bagaimana manusia atau masyarakat dapat mengakses hasil pembangunan dalam dimensi kesehatan, pendidikan dan standar hidup layak.

Salah satu negara yang mengalami peningkatan pembangunan manusia adalah Indonesia. Badan Pusat Statistik (BPS) mencatat bahwa IPM Indonesia pada tahun 2015 mencapai 69,55, sehingga IPM Indonesia meningkat sebesar 0,94 dari tahun sebelumnya (BPS, 2016). Sumatera Utara adalah salah satu provinsi di Indonesia yang memiliki jumlah penduduk terbesar keempat setelah Provinsi Jawa Barat, Jawa Timur dan Jawa Tengah serta mengalami kemajuan dalam pembangunan manusia pada tahun 2015. Hal tersebut ditandai dengan terus meningkatnya IPM Provinsi Sumatera Utara

pada tahun 2015 mencapai 69,51, tumbuh sebesar 0,93 dibandingkan tahun 2014. Provinsi Sumatera Utara mengalami peningkatan IPM dan menduduki peringkat ke sepuluh di tingkat nasional, namun pada tahun 2015 peringkat Provinsi Sumatera Utara masih di bawah dari provinsi lain seperti Bali, Riau, Kalimantan Timur, Kepulauan Riau, dan Sulawesi Utara. IPM Provinsi Sumatera Utara juga masih berada di bawah angka nasional yaitu sebesar 69,55 dan terdapat beberapa Kabupaten/Kota di Provinsi Sumatera Utara yang berstatus rendah karena memiliki nilai IPM rendah. Beberapa Kabupaten/Kota yang memiliki IPM rendah adalah Kabupaten Nias Barat, Kabupaten Nias Selatan, Kabupaten Nias dan Kabupaten Nias Utara.

Perhitungan IPM didasarkan pada tiga dimensi diantaranya dimensi kesehatan, pendidikan dan standar hidup layak. Dimensi kesehatan diwakili oleh indikator angka harapan hidup saat lahir (AHH). AHH di Provinsi Sumatera Utara masih berada di peringkat ke-24 yang termasuk urutan cukup rendah dan paling rendah apabila dibandingkan dengan provinsi lain di Pulau Sumatera. Dimensi pendidikan diwakili oleh indikator Harapan Lama Sekolah (HLS) dan Rata-rata Lama Sekolah (RLS). Berdasarkan HLS pada tahun 2015 Provinsi Sumatera Utara berada di peringkat ke-13 secara nasional, sedangkan apabila dilihat dari RLS maka Provinsi Sumatera Utara berada pada peringkat kelima secara nasional. Dimensi terakhir yaitu standar hidup layak, yang diukur melalui indikator rata-rata pengeluaran per kapita yang disesuaikan. Rata-rata pengeluaran per kapita (PPP) Provinsi Sumatera Utara pada tahun 2015 masih di bawah rata-rata pengeluaran per kapita nasional dan Provinsi Sumatera Utara berada di peringkat ke-17 secara nasional (BPS Sumatera Utara, 2016). Ketiga dimensi yang digunakan dalam mengukur IPM di Provinsi Sumatera Utara masih membutuhkan peningkatan, sehingga perlu dilakukan penelitian untuk mengetahui faktor-faktor yang berpengaruh

terhadap IPM di Provinsi Sumatera Utara. Hal ini bertujuan agar para penyelenggara pemerintah provinsi setempat dapat memperhatikan dan mengupayakan program-program pembangunan manusia untuk meningkatkan nilai IPM sebagai usaha pemerintah dalam perbaikan kesejahteraan masyarakat.

Salah satu metode yang digunakan untuk mengetahui faktor-faktor yang berpengaruh terhadap IPM di Sumatera Utara adalah analisis regresi. Analisis regresi adalah suatu metode statistika yang digunakan untuk mengetahui hubungan antara variabel respon dan variabel prediktor. Ada tiga pendekatan dalam metode analisis regresi untuk mengestimasi kurva regresi yaitu pendekatan parametrik, pendekatan nonparametrik dan pendekatan semiparametrik (Budiantara, 2007). Regresi parametrik digunakan apabila kurva regresi mengikuti pola tertentu atau membentuk pola data yang jelas seperti linier, kuadrat, dan kubik. Regresi nonparametrik digunakan apabila kurva regresi tidak diketahui polanya, sedangkan regresi semiparametrik digunakan apabila terdapat komponen parametrik dan nonparametrik. Penggunaan metode regresi nonparametrik *Spline* memiliki kelebihan dikarenakan *Spline* merupakan model yang mempunyai interpretasi statistik dan visual yang sangat khusus serta sangat baik (Budiantara, 2006). Kelebihan lain dari *Spline* adalah model cenderung mencari sendiri estimasi data kemanapun pola data tersebut bergerak. Kelebihan ini terjadi karena *Spline* terdapat titik-titik knot, yaitu titik perpaduan bersama yang menunjukkan terjadinya perubahan pola data (Budiantara, 2009).

Penelitian ini menggunakan variabel penelitian berdasarkan tiga dimensi dasar untuk mengukur IPM yaitu dimensi kesehatan, pendidikan dan standar hidup layak. Dimensi kesehatan dapat didekati dengan variabel angka harapan hidup. Dimensi pengetahuan dapat didekati menggunakan variabel angka partisipasi murni SD, angka partisipasi murni SMP, angka partisipasi murni SMA,

sedangkan dimensi standar hidup layak menggunakan variabel tingkat partisipasi angkatan kerja dan tingkat pengangguran terbuka. Berdasarkan data yang digunakan dalam penelitian ini, apabila dilihat dari kurva regresi antara nilai IPM sebagai variabel respon dengan faktor-faktor yang memengaruhi nilai IPM sebagai variabel prediktor tidak membentuk pola tertentu. Oleh karena itu, metode analisis yang digunakan dalam penelitian ini adalah regresi nonparametrik Spline. Penelitian sebelumnya tentang IPM telah dilakukan oleh beberapa peneliti diantaranya Ardhanacitri (2013) menggunakan metode *Geographically Weighted Regression* yang menghasilkan kesimpulan yaitu pemetaan pendidikan berdasarkan variabel yang sejalan linier dengan variabel respon yaitu IPM terlihat bahwa terdapat beberapa kabupaten dan kota yang memiliki variabel prediktor yang sama dan terbagi dalam empat kelompok. Meilliana (2013) menggunakan metode regresi panel yang menghasilkan kesimpulan yaitu IPM di Jawa Timur dipengaruhi rasio siswa terhadap guru, angka partisipasi SMP/MTs, jumlah sarana kesehatan, rumah tangga dengan akses air bersih, kepadatan penduduk, tingkat partisipasi angkatan kerja dan PDRB per kapita. Retno (2014) pernah melakukan penelitian tentang IPM dengan menggunakan metode regresi semiparametrik Spline menghasilkan kesimpulan bahwa faktor-faktor yang memengaruhi IPM di Jawa Timur yaitu angka kematian bayi, pertumbuhan ekonomi, tingkat pengangguran terbuka dan tingkat partisipasi angkatan kerja. Penelitian IPM juga pernah dilakukan oleh Yanthi (2016) dengan menggunakan metode regresi nonparametrik Spline yang menghasilkan kesimpulan bahwa faktor-faktor yang memengaruhi IPM di Jawa Tengah yaitu tingkat partisipasi angkatan kerja, rasio sekolah-siswa, kepadatan penduduk, angka kesakitan, PDRB per satu juta. Penelitian tentang IPM di Provinsi Sumatera Utara menggunakan metode regresi

nonparametrik *Spline* masih belum ada yang melakukannya hingga saat ini.

1.2 Rumusan Masalah

Bentuk keberhasilan pembangunan suatu wilayah dapat dilihat dari tingkat pembangunan manusia pada wilayah tersebut. Pembangunan manusia merupakan suatu bentuk investasi modal manusia dalam usaha ikut serta dalam pembangunan nasional. Oleh karena itu, dibutuhkan keseriusan pemerintah dalam menangani masalah peningkatan pembangunan manusia. Salah satu permasalahan pembangunan manusia di Provinsi Sumatera Utara adalah capaian kinerja IPM yang memiliki kecenderungan meningkat secara absolut. Namun peningkatan tersebut ternyata tidak cukup kuat untuk mengangkat posisi relatif IPM di Sumatera Utara ke level yang diharapkan. Capaian kinerja IPM Sumatera Utara diharapkan dapat menjadi lebih baik dengan mengingat berbagai dimensi pembangunan daerah yang masih berada di bawah provinsi lain seperti Bali, Riau, Kalimantan Timur, Kepulauan Riau, dan Sulawesi Utara pada tahun 2015. Selain itu nilai IPM Sumatera Utara juga masih di bawah angka nasional.

Berdasarkan uraian singkat ini, diperoleh rumusan masalah dalam penelitian adalah bagaimana karakteristik dan pemodelan Indeks Pembangunan Manusia di Provinsi Sumatera Utara tahun 2015 beserta faktor-faktor yang memengaruhinya dengan regresi nonparametrik *Spline*. Metode ini dipilih karena mampu memodelkan data yang memiliki perubahan pola perilaku pada sub-sub interval tertentu.

1.3 Tujuan

Tujuan yang ingin dicapai dari penelitian ini adalah sebagai berikut.

1. Mendeskripsikan karakteristik dari Indeks Pembangunan Manusia di Provinsi Sumatera Utara dan faktor-faktor yang memengaruhinya.
2. Memodelkan faktor-faktor yang memengaruhi Indeks Pembangunan Manusia di Provinsi Sumatera Utara dengan regresi nonparametrik Spline.

1.4 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat bagi beberapa pihak berikut.

1. Memberikan informasi tambahan bagi pemerintah Sumatera Utara mengenai faktor-faktor yang memengaruhi Indeks Pembangunan Manusia sehingga pemerintah dapat merencanakan program peningkatan nilai Indeks Pembangunan Manusia di Sumatera Utara.
2. Mampu menerapkan ilmu dan teori Statistika tentang metode regresi nonparametrik Spline khususnya untuk pemodelan Indeks Pembangunan Manusia di Provinsi Sumatera Utara.
3. Penelitian ini mampu dijadikan sumber informasi atau acuan dalam penelitian selanjutnya yang akan dilakukan.

1.5 Batasan Masalah

Batasan masalah yang digunakan pada penelitian ini adalah sebagai berikut.

1. Data yang digunakan adalah data sekunder mengenai Indeks Pembangunan Manusia di Provinsi Sumatera Utara tahun 2015.
2. Metode yang digunakan dalam pemodelan adalah Spline Truncated Linear dengan satu knot, dua knot, tiga knot dan kombinasi knot.
3. Pemilihan titik knot optimal menggunakan metode *Generalized Cross Validation* (GCV).

BAB II TINJAUAN PUSTAKA

2.1 Statistika Deskriptif

Statistika deskriptif merupakan suatu metode yang berkaitan dengan pengumpulan dan penyajian data sehingga memberikan informasi yang berguna (Walpole, 1995). Statistika deskriptif memuat ukuran pemusatan data dan ukuran penyebaran data. Ukuran pemusatan data yang paling sering digunakan adalah rata-rata, sedangkan untuk ukuran penyebaran data yang paling sering digunakan adalah varians. Rata-rata dari data x_1, x_2, \dots, x_n adalah sebagai berikut (Walpole, 1995).

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

dimana \bar{x} adalah rata-rata dan n menunjukkan banyak data. Sementara standar deviasi diberikan oleh.

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

2.2 Analisis Regresi

Analisis regresi merupakan sebuah metode statistika yang memberikan penjelasan tentang pola hubungan (model) antara dua variabel atau lebih (Draper dan Smith, 1992). Model regresi linear berganda diberikan oleh persamaan 2.3.

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + \varepsilon_i, \quad i=1,2,\dots,n \quad (2.3)$$

Berdasarkan model regresi dalam persamaan 2.3, y_i merupakan variabel respon dan x merupakan variabel prediktor. Analisis regresi terdiri dari tiga pendekatan dalam

mengestimasi model yaitu regresi parametrik, semiparametrik dan nonparametrik. Berdasarkan pendekatan regresi parametrik, terdapat asumsi yang sangat kuat dan kaku yaitu bentuk kurva regresi diketahui misalnya linier, kuadratik, kubik, polinomial derajat p , dan lain-lain, sedangkan dalam regresi nonparametrik bentuk kurva regresi tidak diketahui (Budiantara, 2009).

2.3 Regresi Nonparametrik Spline

Regresi nonparametrik merupakan sebuah metode regresi yang tidak terikat asumsi bentuk kurva regresi tertentu (Budiantara, 2001). Pendekatan regresi nonparametrik digunakan untuk menyelesaikan pola data antara variabel respon dengan variabel prediktor tidak membentuk pola tertentu. Fungsi regresi diasumsikan *smooth* (mulus) yang termuat dalam suatu ruang fungsi tertentu. Berdasarkan pendekatan nonparametrik, data diharapkan mencari sendiri bentuk estimasinya tanpa dipengaruhi oleh faktor subjektivitas dari peneliti, dengan demikian pendekatan regresi nonparametrik memiliki fleksibilitas yang tinggi (Eubank, 1988). Bentuk model regresi nonparametrik secara umum adalah sebagai berikut.

$$y_i = f(x_i) + \varepsilon_i, \quad i=1,2,\dots,n \quad (2.4)$$

Berdasarkan persamaan 2.4, y_i adalah variabel respon, fungsi f adalah kurva regresi dengan x_i sebagai variabel prediktor ke- i dan ε_i merupakan *error random* ke- i yang diasumsikan identik, independen dan berdistribusi Normal (IIDN($0, \sigma^2$)). Menurut Wahda (1990) terdapat beberapa pendekatan regresi nonparametrik diantaranya adalah Spline. Pendekatan regresi nonparametrik dan semiparametrik Spline merupakan model yang mempunyai interpretasi statistik dan visual yang sangat khusus dan sangat baik. Spline memiliki kemampuan yang sangat baik untuk menangani data yang perilakunya berubah-ubah pada sub-sub interval tertentu (Budiantara, 2009). Apabila kurva regresi f didekati dengan

fungsi Spline berorde m dengan titik-titik knot K_1, K_2, \dots, K_r yang diberikan oleh persamaan sebagai berikut.

$$f(x_i) = \sum_{h=0}^m \gamma_h x_i^h + \sum_{k=1}^r \gamma_{m+k} (x_i - K_k)_+^m, \quad k=1, 2, \dots, r \quad (2.5)$$

Berdasarkan persamaan 2.5, orde m merupakan derajat polinomial dan K_k merupakan titik knot atau titik yang menunjukkan perubahan pola data. Apabila persamaan (2.5) disubstitusikan ke persamaan (2.4) maka akan diperoleh persamaan regresi nonparametrik Spline sebagai berikut.

$$y_i = \sum_{h=0}^m \gamma_h x_i^h + \sum_{k=1}^r \gamma_{m+k} (x_i - K_k)_+^m + \varepsilon_i, \quad i=1, 2, \dots, n \quad (2.6)$$

Fungsi $(x_i - K_k)_+^m$ adalah fungsi potongan yang diberikan oleh

$$(x_i - K_k)_+^m = \begin{cases} (x_i - K_k)_+^m, & x_i \geq K_k \\ 0 & , \quad x_i < K_k \end{cases} \quad (2.7)$$

2.4 Pemilihan Titik Knot Optimal

Titik knot merupakan titik perpaduan bersama yang menunjukkan perubahan pola perilaku kurva fungsi Spline pada selang yang berbeda (Hardle, 1990). Wahba (1990) mengungkapkan bahwa suatu metode yang digunakan untuk memilih titik knot yaitu dengan metode *Generalized Cross Validation* (GCV). Menurut Budiantara (2006), titik-titik knot optimal diperoleh dari nilai GCV paling minimum. Metode GCV secara umum dituliskan dalam persamaan 2.8 sebagai berikut.

$$GCV(K_1, K_2, \dots, K_r) = \frac{MSE(K_1, K_2, \dots, K_r)}{[n^{-1} \text{trace}(\mathbf{I} - \mathbf{A}(K_1, K_2, \dots, K_r))]^2} \quad (2.8)$$

dengan

K_1, K_2, \dots, K_r = titik knot yang pertama hingga titik knot ke- r

I = matriks identitas

n = banyak pengamatan

$$A(K_1, K_2, \dots, K_r) = X(X'X)^{-1}X'$$

$$MSE(K_1, K_2, \dots, K_r) = n^{-1}(y - \hat{y})'(y - \hat{y}), \text{ dengan}$$

$$\hat{y} = A(K_1, K_2, \dots, K_r) y$$

2.5 Kriteria Pemilihan Model Terbaik

Salah satu kriteria yang digunakan dalam menentukan model regresi terbaik yaitu menggunakan koefisien determinasi (R^2). Koefisien determinasi adalah proporsi keragaman atau variansi total nilai variabel y yang dapat dijelaskan oleh variabel x (Draper dan Smith, 1992). Semakin tinggi nilai R^2 yang dihasilkan dari suatu model, maka semakin baik pula variabel prediktor dalam model regresi menjelaskan variabilitas variabel respon. Berikut ini adalah formula untuk menentukan nilai R^2 .

$$R^2 = \frac{y'X'y - n\bar{y}^2}{y'y - n\bar{y}^2} \quad (2.9)$$

2.6 Pengujian Parameter Model

Pengujian parameter model dilakukan untuk mengetahui apakah variabel memberikan pengaruh signifikan dalam model. Pengujian ini terdiri dari pengujian secara serentak dan pengujian secara parsial.

2.6.1 Pengujian Secara Serentak

Uji serentak merupakan uji signifikansi model secara keseluruhan atau untuk mengetahui apakah semua variabel prediktor yang dimasukkan dalam model memberikan pengaruh secara bersama-sama atau tidak.

Berikut ini adalah hipotesis yang digunakan dalam uji serentak berdasarkan model regresi nonparametrik Spline pada persamaan (2.6).

$$H_0 : \gamma_1 = \gamma_2 = \dots = \gamma_{m+r} = 0$$

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

dimana $m+r$ merupakan banyaknya parameter dalam model regresi nonparametrik Spline kecuali γ_0 . Statistik uji menggunakan uji F sebagai berikut.

$$F_{hitung} = \frac{MS_{regresi}}{MS_{residual}} \quad (2.10)$$

dengan

$$MS_{regresi} = \frac{SS_{regresi}}{df_{regresi}} = \frac{\gamma' x' y - n y^2}{m+r} \quad \text{dan} \quad MS_{residual} = \frac{SS_{residual}}{df_{residual}} = \frac{y' y - \gamma' x' y}{n - (m+r) - 1}$$

Daerah penolakan H_0 adalah apabila nilai F_{hitung} lebih besar daripada $F_{\alpha, (m+r, n-(m+r)-1)}$ atau $p\text{-value} < \alpha$. Apabila keputusan tolak H_0 maka dapat disimpulkan minimal terdapat satu parameter pada model regresi Spline yang signifikan.

2.6.2 Pengujian Secara Parsial

Apabila dalam pengujian secara serentak didapatkan kesimpulan minimal terdapat satu parameter pada model regresi Spline yang signifikan maka perlu dilanjutkan uji parsial. Pengujian secara parsial dilakukan untuk mengetahui parameter mana yang signifikan dan tidak signifikan dalam model. Berikut ini adalah hipotesis yang digunakan untuk menguji signifikansi model secara parsial.

$$H_0 : \gamma_j = 0$$

$$H_1 : \gamma_j \neq 0, j = 1, 2, \dots, m + r$$

Statistik uji yang digunakan adalah uji t sebagai berikut.

$$t_{hitung} = \frac{\hat{\gamma}_j}{se(\hat{\gamma}_j)}, j = 1, 2, \dots, m+r \quad (2.11)$$

dengan

$\hat{\gamma}_j$: penaksir parameter γ_j

$se(\hat{\gamma}_j)$: standart error dari $\hat{\gamma}_j$

Nilai $se(\hat{\gamma}_j)$ didapatkan dari $\sqrt{diag(var(\hat{\gamma}_j))}$. $Var(\hat{\gamma}_j)$ merupakan elemen diagonal ke- j dari matriks.

$$\begin{aligned} var(\gamma) &= var[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}] \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' var(y)[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']' \\ &= \sigma^2 (\mathbf{X}'\mathbf{X})^{-1} \end{aligned} \quad (2.12)$$

Daerah penolakan yang digunakan adalah tolak H_0 apabila $|t_{hitung}|$ lebih besar daripada $t_{(\frac{\alpha}{2}, n-(m+r)-1)}$ atau $p\text{-value} < \alpha$.

Kesimpulan yang diperoleh apabila tolak H_0 adalah parameter berpengaruh signifikan terhadap model (Draper dan Smith, 1992).

2.7 Pengujian Asumsi Residual

Apabila model terbaik dari regresi Spline telah didapatkan, maka perlu dilakukan pengujian asumsi residual untuk mengetahui apakah residual yang dihasilkan telah memenuhi asumsi yaitu identik, independen dan berdistribusi Normal.

2.7.1 Uji Asumsi Identik

Pengujian ini dilakukan untuk mengetahui varians residual model regresi. Apabila varians antar residual tidak homogen maka terjadi heteroskedastisitas. Heteroskedastisitas dapat dideteksi secara visual dan uji *Glejser*. Cara untuk mendeteksi heteroskedastisitas secara visual dapat dilihat dari *plot* antara residual dan estimasi respon (\hat{y}) menunjukkan sebaran data membentuk suatu pola tertentu. Cara kedua untuk mengetahui heteroskedastisitas yaitu dengan melakukan

uji *Glejser*. Uji ini dilakukan dengan meregresikan nilai mutlak residual dengan variabel prediktor.

$$|\varepsilon_i| = f(x_i) + \varepsilon_i \quad (2.13)$$

Apabila variabel prediktor signifikan dalam model maka varians residual sebanding dengan nilai x , hal ini mengindikasikan bahwa residual cenderung tidak homogen atau terjadi heteroskedastisitas.

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

$$H_1 : \text{minimal terdapat satu } \sigma_i^2 \neq \sigma^2, i=1,2,\dots,n$$

Statistik uji yang digunakan adalah sebagai berikut.

$$F_{hitung} = \frac{[\sum_{i=1}^n (|\hat{e}_i| - |\bar{e}|)^2] / (m+r)}{[\sum_{i=1}^n (|e_i| - |\bar{e}_i|)^2] / (n-(m+r)-1)} \quad (2.14)$$

Daerah penolakan yang digunakan adalah tolak H_0 apabila F_{hitung} lebih besar dari F_{tabel} yaitu $F_{\alpha,(m+r),n-(m+r)-1}$ atau $p\text{-value} < \alpha$, dimana nilai m merupakan parameter model *Glejser* dari derajat polinomial, r adalah parameter komponen *truncated* dan n adalah jumlah observasi. Apabila H_0 ditolak maka dapat disimpulkan bahwa minimal terdapat satu $\sigma_i^2 \neq \sigma^2$. Hal tersebut menandakan terdapat kasus heteroskedastisitas sehingga asumsi identik tidak terpenuhi. Untuk mengatasi kasus heteroskedastisitas yaitu dengan transformasi variabel menggunakan *Weighted Least Square* (Gujarati, 1992).

2.7.2 Uji Asumsi Independen

Menurut Agresti (1990), uji independensi digunakan untuk mengetahui hubungan antara dua variabel. Pengujian asumsi independen dilakukan untuk mengetahui apakah korelasi antar residual bernilai nol atau tidak. Asumsi independen terpenuhi apabila tidak terdapat korelasi antar residual atau autokorelasi. Ada beberapa cara untuk

mendeteksi autokorelasi diantaranya adalah dengan melakukan uji *Durbin Watson* (Gujarati, 2004).

Hipotesis untuk uji *Durbin Watson* adalah sebagai berikut.

$H_0 : \rho = 0$ (Tidak terjadi autokorelasi antar residual)

$H_1 : \rho \neq 0$ (Terjadi autokorelasi antar residual)

Statistik uji yang digunakan adalah

$$d_{hitung} = \frac{\sum_{i=2}^n (\hat{e}_i - \hat{e}_{i-1})^2}{\sum_{i=1}^n \hat{e}_i^2} \quad (2.15)$$

Daerah keputusan terbagi menjadi beberapa bagian yaitu:

- Apabila $d_{hitung} < d_{L,\alpha/2}$ atau $(4 - d_{hitung}) < d_{L,\alpha/2}$, maka tolak H_0 .
- Apabila $d_{hitung} > d_{U,\alpha/2}$ atau $(4 - d_{hitung}) > d_{U,\alpha/2}$, maka gagal tolak H_0 .
- Apabila $d_{L,\alpha/2} \leq d_{hitung} \leq d_{U,\alpha/2}$ atau $d_{L,\alpha/2} \leq (4 - d_{hitung}) < d_{U,\alpha/2}$, maka tidak ada keputusan tolak H_0 atau gagal tolak H_0 .

Metode lain yang digunakan untuk mengidentifikasi terdapat korelasi antar residual atau tidak dengan melihat plot *Autocorrelation Function* (ACF) dari residual. Apabila pada plot ACF menunjukkan tidak terdapat nilai autokorelasi yang keluar dari batas signifikansi maka kesimpulan yang diperoleh adalah tidak terjadi autokorelasi antar residual, sehingga asumsi independen terpenuhi. Formula untuk menghitung nilai autokorelasi yang diberikan oleh persamaan 2.16.

$$\rho_l = \frac{\sum_{t=1}^{n-l} (e_t - \bar{e})(e_{t+l} - \bar{e})}{\sum_{t=1}^n (e_t - \bar{e})^2}, \quad l = 1, 2, 3, \dots \quad (2.16)$$

dengan

ρ_l = korelasi antara e_t dan e_{t+l}

l = lag ke- l

Interval konfidensi $(1-\alpha)100\%$ untuk autokorelasi ρ_l diberikan oleh persamaan 2.17.

$$-t_{n-1, \alpha/2} SE(\rho_l) < \rho_l < t_{n-1, \alpha/2} SE(\rho_l) \quad (2.17)$$

dengan

$$SE(\rho_l) = \sqrt{\frac{1 + 2 \sum_{l=1}^{l-1} \rho_l^2}{n}} \quad (2.18)$$

Gagal tolak H_0 atau asumsi independen terpenuhi apabila tidak terdapat $\hat{\rho}_l$ yang keluar dari batas interval konfidensi (Wei, 2006).

2.7.3 Uji Asumsi Distribusi Normal

Uji asumsi distribusi Normal dilakukan untuk mengetahui apakah residual model regresi berdistribusi Normal atau tidak. Secara visual pengujian asumsi distribusi normal bisa dilakukan dengan *normal probability plot residual*. Cara lain dapat dilakukan pengujian *Kolmogorov Smirnov* dengan hipotesis sebagai berikut.

$H_0: F(x) = F_0(x)$ (Residual berdistribusi Normal)

$H_1: F(x) \neq F_0(x)$ (Residual tidak berdistribusi Normal)

Statistik uji yang digunakan dalam uji *Kolmogorov Smirnov* adalah sebagai berikut.

$$D = \max |F_0(x) - S_N(x)| \quad (2.19)$$

dimana $F_0(x)$ merupakan fungsi distribusi frekuensi kumulatif. Sementara $S_N(x)$ adalah fungsi peluang kumulatif yang diobservasi dari suatu sampel *random*.

Daerah penolakan pada uji *Kolmogorov Smirnov* yaitu tolak H_0 apabila $|D| > q_{(1-\alpha)}$ dimana nilai $q_{(1-\alpha)}$ didapatkan dari tabel *Kolmogorov Smirnov* atau tolak H_0 apabila $p - value < \alpha$. Keputusan yang didapatkan apabila tolak H_0 adalah residual tidak berdistribusi Normal. Transformasi *box-cox* digunakan untuk mengatasi apabila asumsi residual berdistribusi Normal tidak terpenuhi.

2.8 Indeks Pembangunan Manusia

Menurut UNDP, IPM mengukur capaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup. IPM merupakan indeks komposit yang dihitung sebagai rata-rata sederhana dari tiga indeks dasar yaitu indeks kesehatan, pendidikan dan standar hidup layak (BPS, 2016). UNDP memperkenalkan penghitungan IPM metode baru dengan beberapa perbedaan yang mendasar dibanding metode lama. Ada dua hal mendasar dalam perubahan metode baru ini yaitu cara penghitungan indeks dan aspek indikator. UNDP melakukan perubahan cara penghitungan indeks. Penghitungan IPM dengan metode lama menggunakan rata-rata aritmatik, sedangkan dengan metode baru menggunakan rata-rata geometrik. Cara penghitungan IPM dengan menggunakan rata-rata geometrik akan cenderung sensitif terhadap ketimpangan. Rata-rata geometrik menuntut keseimbangan ketiga dimensi IPM agar capaian IPM menjadi optimal. Berikut ini merupakan definisi dari komponen pembentukan IPM menggunakan metode baru.

a. Angka Harapan Hidup

Angka Harapan Hidup (AHH) merupakan rata-rata perkiraan banyak tahun yang dapat ditempuh oleh seseorang sejak lahir. AHH mencerminkan derajat kesehatan suatu masyarakat. AHH dihitung dari hasil sensus dan survei kependudukan.

b. Harapan Lama Sekolah

Harapan Lama Sekolah (HLS) adalah lamanya sekolah (dalam tahun) yang diharapkan akan dirasakan oleh anak

- pada umur tertentu di masa mendatang. HLS dihitung untuk penduduk berusia 7 tahun ke atas. HLS dapat digunakan untuk mengetahui kondisi pembangunan sistem pendidikan di berbagai jenjang yang ditunjukkan dalam bentuk lamanya pendidikan (dalam tahun) yang diharapkan dapat dicapai oleh setiap anak.
- c. Rata-rata Lama Sekolah
Rata-rata lama sekolah (RLS) adalah jumlah tahun yang digunakan oleh penduduk dalam menjalani pendidikan formal. Cakupan penduduk yang dihitung dalam penghitungan rata-rata lama sekolah adalah penduduk berusia 25 tahun ke atas.
- d. Pengeluaran per Kapita Disesuaikan
Pengeluaran per Kapita Disesuaikan (PPP) ditentukan dari nilai pengeluaran per kapita dan paritas daya beli. Rata-rata pengeluaran per kapita setahun diperoleh dari Susenas, dihitung dari level provinsi hingga level kabupaten/kota. Perhitungan paritas daya beli pada metode baru menggunakan 96 komoditas dimana 66 komoditas merupakan makanan dan sisanya merupakan komoditas non makanan.

Capaian pembangunan manusia di suatu wilayah pada waktu tertentu dapat dikelompokkan ke dalam empat kelompok. Pengelompokkan ini disajikan dalam Tabel 2.1 yang bertujuan untuk mengorganisasikan wilayah-wilayah menjadi kelompok-kelompok yang sama dalam hal pembangunan manusia (BPS Sumatera Utara, 2016).

Tabel 2.1 Penentuan Status IPM Berdasarkan Nilai IPM

Status IPM	Nilai IPM
Rendah	$IPM < 60$
Sedang	$60 \leq IPM < 70$
Tinggi	$70 \leq IPM < 80$
Sangat Tinggi	$IPM \geq 80$

2.9 Kerangka Konsep

Pemilihan variabel penelitian dalam penelitian kali ini berdasarkan penelitian sebelumnya, pakar ekonomi dan informasi dari Badan Pusat Statistik. Berikut ini adalah penjelasan dari pemilihan variabel penelitian.

a. Dimensi Kesehatan

Salah satu dimensi untuk mengukur IPM adalah dimensi kesehatan. Dimensi kesehatan dalam penelitian ini, didekati oleh variabel Angka Harapan Hidup (AHH). AHH adalah indikator yang mengukur *longevity* (panjang umur) dari seseorang di suatu wilayah atau negara. *Longevity* ini bukan hanya upaya perorangan tetapi merupakan upaya masyarakat secara keseluruhan untuk menggunakan sumber daya yang ada sehingga dapat memperpanjang hidupnya, sehingga dapat dikatakan seseorang akan bertahan hidup lebih panjang apabila selalu sehat, atau apabila menderita sakit secepatnya dapat berobat untuk membantu mempercepat kesembuhannya (Suparman, 1986 dalam Bappeda Kabupaten Malang, 2014).

b. Dimensi Pendidikan

Indeks Pembangunan Manusia dipengaruhi oleh pertumbuhan ekonomi. Pendidikan memiliki pengaruh penting terhadap pertumbuhan ekonomi. Tingkat pendidikan seseorang tersebut dipengaruhi oleh tingkat produktivitas barang dan jasa. Tingkat pendidikan yang tinggi akan mempengaruhi kualitas kinerja di perusahaan sehingga diharapkan mampu menghasilkan suatu *output* yang produktif. Upaya untuk meningkatkan kualitas sumber daya manusia melalui pendidikan dengan cara program wajib belajar 12 tahun (BPS,2016). Oleh karena itu, dalam penelitian ini dimensi pendidikan didekati oleh variabel angka partisipasi murni SD, angka partisipasi murni SMP dan angka partisipasi murni SMA.

c. Dimensi Standar Hidup Layak

Dimensi standar hidup layak dalam penelitian ini, didekati oleh variabel Tingkat Partisipasi Angkatan Kerja (TPAK) dan

Tingkat Pengangguran Terbuka (TPT). Tenaga kerja merupakan faktor yang memengaruhi pertumbuhan ekonomi. Peranan tenaga kerja tersebut sebagai salah satu faktor produksi yang akan memengaruhi tinggi rendahnya tingkat pendapatan nasional dari segi kuantitas atau jumlah saja. Sementara itu kita beranggapan bahwa apabila jumlah tenaga kerja yang dipakai dalam usaha produksi meningkat, maka jumlah produksi yang bersangkutan juga meningkat. Dengan kata lain kalau tidak ada peningkatan jumlah tenaga kerja maka jumlah produksi akan tetap. Apabila kualitas dari tenaga kerja itu menjadi lebih baik, maka dapat terjadi bahwa tingkat produksi akan meningkat pula (Irawan dan Suparmoko, 2008).

Todaro (2000) mengatakan bahwa pembangunan manusia memainkan peranan kunci dalam membentuk kemampuan sebuah negara dalam menyerap teknologi modern untuk mengembangkan kapasitasnya agar tercipta kesempatan kerja untuk mengurangi jumlah pengangguran untuk melakukan pembangunan manusia yang berkelanjutan. Upaya mengatasi jumlah pengangguran dan mendapatkan pendapatan yang tinggi maka akan berpengaruh terhadap peningkatan pembangunan manusia melalui peningkatan bagian pengeluaran rumah tangga yang dibelanjakan untuk makanan yang lebih bergizi dan pendidikan yang lebih tinggi, sehingga pengurangan pengangguran dapat terlihat dari jumlah indeks pembangunan manusia yang mengalami peningkatan.

Menurut Sadono Sukirno (2004), efek buruk dari pengangguran adalah mengurangi pendapatan masyarakat yang pada akhirnya mengurangi tingkat kemakmuran dan kesejahteraan yang telah dicapai seseorang. Semakin turunnya kesejahteraan masyarakat karena pengangguran tentunya akan meningkatkan peluang terjebak pada rendahnya Indeks Pembangunan Manusia karena tidak dapat memiliki pendapatan untuk memenuhi kebutuhannya. Apabila pengangguran di suatu negara sangat buruk, kekacauan politik

dan sosial selalu berlaku dan menimbulkan efek yang buruk bagi kesejahteraan masyarakat dan prospek meningkatkan Indeks Pembangunan Manusia dalam jangka menengah sampai jangka panjang.

Teori pertumbuhan ekonomi baru menekankan pentingnya peranan pemerintah terutama dalam meningkatkan pembangunan modal manusia dan mendorong penelitian dan pengembangan untuk meningkatkan produktivitas manusia. Perusahaan akan memperoleh hasil yang lebih banyak dengan memperkerjakan tenaga kerja dengan produktivitas yang tinggi, sehingga perusahaan juga akan bersedia memberikan gaji yang lebih tinggi bagi yang bersangkutan. Seseorang yang memiliki produktivitas yang tinggi akan memperoleh kesejahteraan yang lebih baik, yang diperlihatkan melalui peningkatan pendapatan maupun konsumsinya. Upaya penyerapan tenaga kerja yang semakin banyak menyebabkan berkurangnya tingkat pengangguran yang selama ini terjadi.

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diambil dari Badan Pusat Statistik (BPS) Provinsi Sumatera Utara dengan judul publikasi Provinsi Sumatera Utara Dalam Angka 2016 dan Beberapa Indikator Penting Provinsi Sumatera Utara 2015/2016. Data yang digunakan adalah data pada tahun 2015 dengan banyaknya observasi sebesar 33 Kabupaten/Kota di Sumatera Utara.

3.2 Variabel Penelitian

Variabel penelitian terdiri dari variabel respon dan variabel prediktor. Variabel respon dalam penelitian ini adalah Indeks Pembangunan Manusia di Sumatera Utara pada tahun 2015. Variabel prediktor yang memengaruhi Indeks Pembangunan Manusia di Sumatera Utara dalam penelitian ini diperoleh dari penelitian-penelitian sebelumnya yang dapat diuraikan dalam Tabel 3.1.

Tabel 3.1 Variabel Penelitian

Variabel	Keterangan
y	Indeks Pembangunan Manusia
x_1	Angka Partisipasi Murni SD
x_2	Angka Partisipasi Murni SMP
x_3	Angka Partisipasi Murni SMA
x_4	Tingkat Partisipasi Angkatan Kerja
x_5	Tingkat Pengangguran Terbuka
x_6	Angka Harapan Hidup

Berikut ini adalah keterangan mengenai variabel penelitian yang digunakan.

- a. Variabel y adalah variabel respon yang menyatakan Indeks Pembangunan Manusia (IPM). IPM dapat digunakan

sebagai ukuran kualitas hidup yang dibangun melalui tiga dimensi dasar. Dimensi tersebut mencakup kesehatan, pendidikan, dan standar hidup yang layak. Ketiga dimensi tersebut memiliki pengertian sangat luas karena terkait banyak faktor. Pengukuran dimensi kesehatan dapat menggunakan angka harapan hidup waktu lahir, sedangkan untuk mengukur dimensi pendidikan digunakan gabungan indikator harapan lama sekolah dan rata-rata lama sekolah. Dimensi standar hidup layak menggambarkan tingkat kesejahteraan yang dinikmati oleh penduduk sebagai dampak semakin membaiknya ekonomi apabila dilihat dalam cakupan lebih luas. UNDP mengukur standar hidup layak menggunakan pengeluaran per kapita yang disesuaikan (BPS, 2016).

- b. Variabel x_1 merupakan variabel prediktor yang menyatakan Angka Partisipasi Murni (APM) SD. APM SD adalah proporsi anak sekolah pada satu kelompok usia 7-12 tahun yang bersekolah pada jenjang yang sesuai dengan kelompok usianya. APM SD meliputi Sekolah Dasar, Madrasah Ibtidaiyah dan sederajat (BPS, 2016). Perhitungan APM SD didapatkan dari:

$$\text{APM SD} = \frac{\text{Jumlah murid SD usia 7 - 12 tahun}}{\text{Jumlah penduduk usia 7 - 12 tahun}} \times 100$$

- c. Variabel x_2 merupakan variabel prediktor yang menyatakan Angka Partisipasi Murni (APM) SMP. APM SMP merupakan proporsi anak sekolah pada satu kelompok usia 13-15 tahun yang bersekolah pada jenjang yang sesuai dengan kelompok usianya (BPS, 2016). APM SMP meliputi SMP umum, Madrasah Tsanawiyah, SMP kejuruan dan sederajat. APM SMP diperoleh dari:

$$\text{APM SMP} = \frac{\text{Jumlah murid SMP usia 13 - 15 tahun}}{\text{Jumlah penduduk usia 13 - 15 tahun}} \times 100$$

- d. Variabel x_3 merupakan variabel prediktor yang menyatakan Angka Partisipasi Murni (APM) SMA. APM SMA merupakan proporsi anak sekolah pada satu kelompok usia 16-18 tahun yang bersekolah pada jenjang yang sesuai dengan kelompok usianya (BPS, 2016). APM SMA meliputi SMA, SMK, Madrasah Aliyah dan sederajat. APM SMA didapatkan dari:

$$\text{APM SMA} = \frac{\text{Jumlah murid SMA usia 16-18 tahun}}{\text{Jumlah penduduk usia 16-18 tahun}} \times 100$$

- e. Variabel x_4 merupakan variabel prediktor yang menyatakan Tingkat Partisipasi Angkatan Kerja (TPAK). TPAK adalah angka yang menunjukkan persentase angkatan kerja terhadap penduduk usia kerja (Mantra, 2003). TPAK merupakan indikator yang mampu menggambarkan sejauh mana peran angkatan kerja di suatu daerah. Semakin tinggi nilai TPAK maka semakin besar pula keterlibatan penduduk usia kerja dalam pasar kerja (BPS Sumatera Utara, 2015). TPAK secara umum dapat dirumuskan:

$$\text{TPAK} = \frac{\text{Angkatan Kerja}}{\text{Penduduk Usia Kerja}} \times 100$$

- f. Variabel x_5 adalah variabel prediktor yang menyatakan Tingkat Pengangguran Terbuka (TPT). TPT merupakan persentase jumlah pengangguran terhadap jumlah angkatan kerja. TPT secara umum dapat dirumuskan sebagai berikut:

$$\text{TPT} = \frac{\text{Jumlah Pengangguran}}{\text{Jumlah Angkatan Kerja}} \times 100$$

- g. Variabel x_6 adalah variabel prediktor yang menyatakan Angka Harapan Hidup (AHH). Untuk menghitung AHH (e_0), idealnya dihitung berdasarkan Angka Kematian Menurut Umur (*Age Specific Death Rate/ASDR*) yang

datanya diperoleh dari catatan registrasi kematian secara bertahun-tahun sehingga dimungkinkan dibuat Tabel Kematian. Tetapi karena sistem registrasi penduduk di Indonesia belum berjalan dengan baik maka untuk menghitung Angka Harapan Hidup digunakan cara tidak langsung dengan program *Mortpak Lite* (BPS, 2016).

3.3 Struktur Data

Berikut ini adalah struktur data dari variabel respon dan variabel prediktor yang digunakan dalam penelitian ini diuraikan dalam Tabel 3.2.

Tabel 3.2 Struktur Data Penelitian

Kab./Kota	y	x_1	x_2	x_3	x_4	x_5	x_6
1	y_1	$x_{1,1}$	$x_{2,1}$	$x_{3,1}$	$x_{4,1}$	$x_{5,1}$	$x_{6,1}$
2	y_2	$x_{1,2}$	$x_{2,2}$	$x_{3,2}$	$x_{4,2}$	$x_{5,2}$	$x_{6,2}$
3	y_3	$x_{1,3}$	$x_{2,3}$	$x_{3,3}$	$x_{4,3}$	$x_{5,3}$	$x_{6,3}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
33	y_{33}	$x_{1,33}$	$x_{2,33}$	$x_{3,33}$	$x_{4,33}$	$x_{5,33}$	$x_{6,33}$

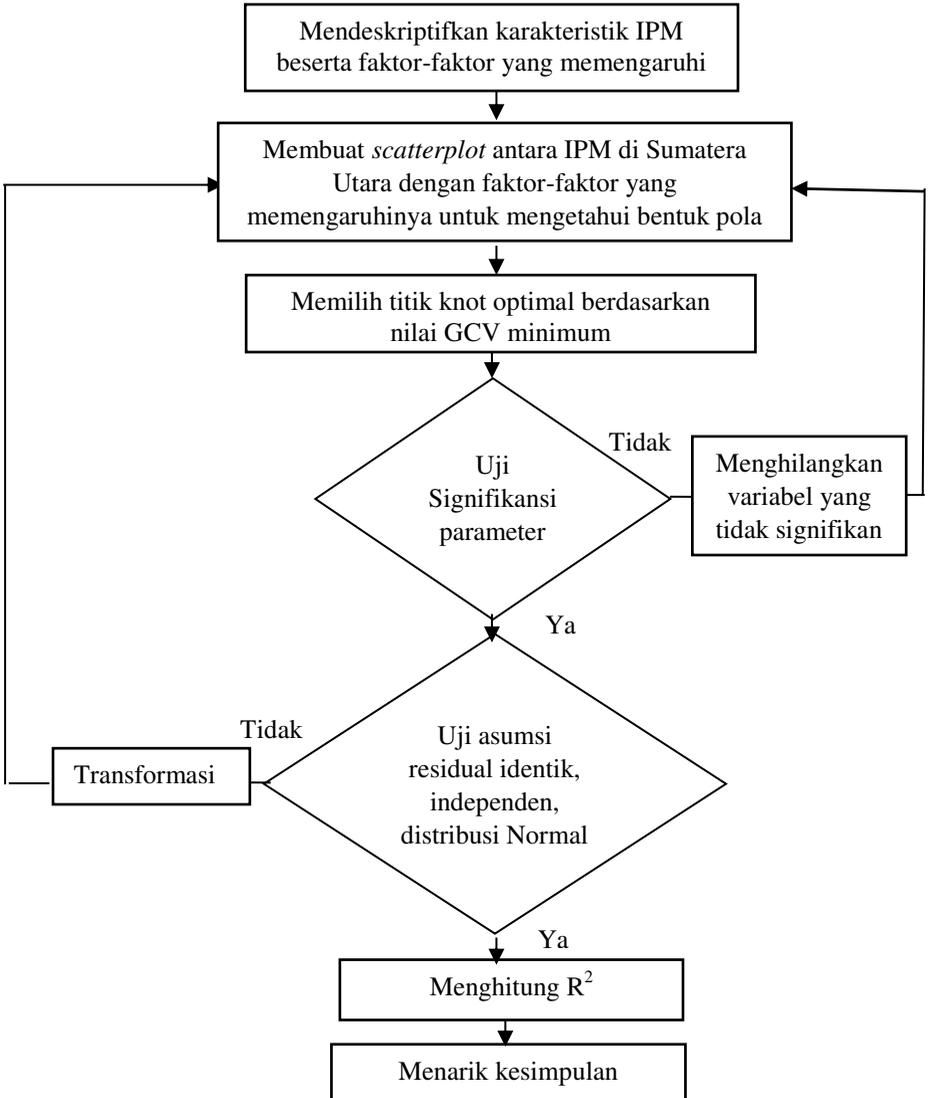
3.4 Langkah Analisis

Langkah-langkah analisis yang dilakukan dalam penelitian ini adalah sebagai berikut.

1. Mendeskripsikan karakteristik Indeks Pembangunan Manusia di Sumatera Utara beserta faktor-faktor yang memengaruhinya.
2. Memodelkan faktor-faktor yang memengaruhi Indeks Pembangunan Manusia menggunakan regresi nonparametrik Spline.
 - a. Membuat *scatterplot* antara Indeks Pembangunan Manusia di Sumatera Utara dengan variabel-variabel prediktornya untuk mengetahui bentuk pola data. Apabila bentuk pola data tidak mengikuti pola tertentu maka metode yang digunakan yaitu regresi nonparametrik Spline.

- b. Memodelkan data menggunakan Spline Truncated Linear dengan satu knot, dua knot, tiga knot dan kombinasi knot.
- c. Mendapatkan model regresi nonparametrik Spline dengan titik knot optimal berdasarkan nilai *Generalized Cross Validation* (GCV) yang paling minimum.
- d. Melakukan pengujian signifikansi parameter regresi Spline secara serentak dan parsial.
- e. Melakukan uji asumsi residual identik, independen dan berdistribusi Normal (IIDN) berdasarkan model regresi Spline. Apabila residual model regresi Spline tidak memenuhi asumsi IIDN maka harus dilakukan transformasi data. Setelah melakukan transformasi data, memulai analisis kembali dari langkah (a).
- f. Menghitung koefisien determinasi (R^2).
- g. Melakukan interpretasi model.
- h. Menarik kesimpulan dan saran.

Gambar 3.1 menyajikan langkah-langkah analisis yang dibentuk dalam diagram alir.



Gambar 3.1 Diagram Alir

BAB IV ANALISIS DAN PEMBAHASAN

Bab ini berisi pembahasan mengenai karakteristik data, pola hubungan Indeks Pembangunan Manusia (IPM) di Provinsi Sumatera Utara dan 6 variabel yang diduga memengaruhinya, pemodelan IPM Provinsi Sumatera Utara menggunakan regresi nonparametrik Spline, pemilihan titik knot optimal untuk mendapatkan model terbaik, uji signifikansi parameter model secara serentak dan parsial serta uji asumsi residual.

4.1 Karakteristik Data Indeks Pembangunan Manusia

Indeks Pembangunan Manusia (IPM) dibentuk oleh tiga dimensi dasar yaitu dimensi kesehatan, pendidikan dan standar hidup layak. IPM Sumatera Utara mengalami peningkatan setiap tahunnya. Hal ini dibuktikan dengan data pada tahun 2015, Sumatera Utara memiliki IPM sebesar 69,51 yang mengalami peningkatan sebesar 0,93 persen dibandingkan tahun sebelumnya. Berdasarkan BPS, status IPM dapat dikategorikan menjadi 4 yaitu rendah ($IPM < 60$), sedang ($60 \leq IPM < 70$), tinggi ($70 \leq IPM < 80$) dan sangat tinggi ($IPM \geq 80$).

Tabel 4.1 Jumlah Kabupaten/Kota Berdasarkan Status IPM

Status IPM	Jumlah Kab./Kota
Rendah	4
Sedang	17
Tinggi	12
Sangat Tinggi	-

Berdasarkan Tabel 4.1, keempat kategori IPM tersebar pada setiap Kabupaten/Kota. Terdapat 4 kabupaten yang terkategori sebagai wilayah yang memiliki nilai IPM rendah antara lain Kabupaten Nias Barat, Kabupaten Nias Selatan, Kabupaten Nias dan Kabupaten Nias Utara. Menurut BPS Sumatera Utara, keempat Kabupaten yang tergolong dalam status IPM rendah merupakan kawasan Pantai Barat, dimana kawasan tersebut masih mengalami kesenjangan pembangunan manusia, sehingga diharap-

kan pemerintah dapat membuat program-program untuk meningkatkan IPM dari keempat wilayah tersebut.

Wilayah di Provinsi Sumatera Utara dengan status IPM sedang ($60 \leq \text{IPM} < 70$) ada 17 Kabupaten/Kota. Wilayah tersebut adalah sebagai berikut.

Tabel 4.2 Kabupaten/Kota yang Terkategorikan dalam Status IPM Sedang

No.	Kab./Kota	IPM
1	Kab. Mandailing Natal	63,99
2	Kab. Tapanuli Selatan	67,63
3	Kab. Tapanuli Tengah	67,06
4	Kab. Asahan	68,40
5	Kab. Dairi	69,00
6	Kab. Langkat	68,53
7	Kab. Humbang Hasundutan	66,03
8	Kab. Pakpak Bharat	65,53
9	Kab. Samosir	68,43
10	Kab. Serdang Bedagai	68,01
11	Kab. Batu Bara	66,02
12	Kab. Padang Lawas Utara	67,35
13	Kab. Padang Lawas	65,99
14	Kab. Labuhanbatu Selatan	69,67
15	Kab. Labuanbatu Utara	69,69
16	Kota Tanjungbalai	66,74
17	Kota Gunungsitoli	66,41

Menurut informasi dari BPS Sumatera Utara, sebagian besar wilayah yang tergolong dalam status IPM sedang merupakan kawasan Pantai Timur. Kawasan tersebut merupakan kawasan yang mengalami perkembangan cukup pesat karena persyaratan infrastruktur relatif lebih lengkap dibandingkan dengan kawasan Pantai Barat. Selanjutnya berdasarkan Tabel 4.1, terdapat 12 Kabupaten/Kota yang berstatus IPM tinggi ($70 \leq \text{IPM} < 80$). Menurut BPS Sumatera Utara, sebagian besar Kabupaten/Kota yang tergolong dalam status IPM tinggi adalah kawasan dataran tinggi. Kawasan dataran tinggi merupakan kawasan yang memiliki sumber daya alam yang memadai untuk dikelola

pemerintah sehingga berdampak baik bagi kesejahteraan masyarakat wilayah tersebut.

Tabel 4.3 Kabupaten/Kota yang Terkategorikan dalam Status IPM Tinggi

No.	Kab./Kota	IPM
1	Kab. Tapanuli Utara	71,32
2	Kab. Toba Samosir	73,40
3	Kab. Labuhan Batu	70,23
4	Kab. Simalungun	71,24
5	Kab. Karo	72,69
6	Kab. Deli Serdang	72,79
7	Kota Sibolga	71,64
8	Kota Pematangsiantar	76,34
9	Kota Tebing Tinggi	72,81
10	Kota Medan	78,87
11	Kota Binjai	73,81
12	Kota Padangsidimpuan	72,80

Berdasarkan data IPM Sumatera Utara tahun 2015, masih belum terdapat Kabupaten/Kota yang berstatus IPM sangat tinggi ($IPM \geq 80$) sehingga pemerintah Sumatera Utara perlu memerhatikan dan meningkatkan faktor-faktor apa saja yang berpengaruh terhadap IPM.

Statistika deskriptif terdiri dari ukuran pemusatan data dan ukuran penyebaran data. Tabel 4.4 merupakan tabel yang memberikan informasi tentang statistika deskriptif yaitu rata-rata, standar deviasi, nilai minimum dan maksimum.

Tabel 4.4 Statistika Deskriptif dari Variabel Penelitian

Variabel	Rata-rata	Standar Deviasi	Minimum	Maksimum
Y	68,42	4,85	58,25	78,87
X ₁	97,20	2,75	88,96	100,00
X ₂	77,26	5,73	68,51	91,42
X ₃	67,21	8,84	47,55	87,53
X ₄	72,80	9,66	57,72	88,38
X ₅	5,84	3,31	0,40	11,39
X ₆	68,07	2,56	61,58	72,29

Berdasarkan Tabel 4.4, ditunjukkan bahwa variabel respon (y) merupakan variabel IPM yang memiliki rata-rata sebesar 68,42

dengan standar deviasi sebesar 4,85. Nilai minimum dari variabel IPM sebesar 58,25 yang terdapat di Kabupaten Nias Barat, sedangkan nilai maksimum dari variabel IPM adalah 78,87 yang terdapat di Kota Medan. Kabupaten Nias Barat adalah wilayah yang memiliki IPM terendah karena masih rendahnya sumber daya manusia dan infrastruktur yang dimiliki oleh Kabupaten Nias Barat. Hal ini didukung dengan adanya jabatan kepala dinas pekerjaan umum atau kepala badan perencanaan pembangunan yang diisi oleh orang di luar kompetensinya. Kota Medan merupakan wilayah dengan IPM tertinggi karena pertumbuhan ekonomi Kota Medan dari tahun 2010 mengalami peningkatan diatas rata-rata setiap tahunnya. Pertumbuhan ekonomi yang terus meningkat berdampak positif terhadap meningkatnya produktivitas dan nilai tambah sektor-sektor yang menyerap tenaga kerja seperti pertanian, perkebunan, kelautan dan perikanan.

Angka Partisipasi Murni SD (x_1) merupakan salah satu faktor yang diduga berpengaruh terhadap IPM. Angka Partisipasi Murni (APM) SD dapat menggambarkan dimensi pendidikan yang terdapat dalam indikator IPM. Berdasarkan Tabel 4.4, terlihat bahwa variabel APM SD memiliki rata-rata sebesar 97,20 dengan standar deviasi sebesar 2,75. Nilai minimum dari variabel APM SD sebesar 88,96 persen yang terdapat di Kota Medan, sedangkan nilai maksimum dari variabel APM SD adalah 100 persen yang terdapat di Kabupaten Tapanuli Utara. Kota Medan merupakan wilayah yang memiliki APM SD terendah karena fasilitas pendidikan dasar yang diberikan pemerintah belum memadai, sehingga pemerintah Kota Medan memerlukan bantuan dari pihak swasta agar fasilitasnya terpenuhi. Kabupaten Tapanuli Utara memiliki APM SD tertinggi karena penduduk setempat masih memerhatikan pendidikan khususnya program wajib belajar 12 tahun, walaupun sebagian besar penduduk bermata pencaharian di sektor pertanian yang tergolong pendapatannya rendah.

Variabel x_2 merupakan Angka Partisipasi Murni (APM) SMP yang diduga memengaruhi IPM. Variabel ini dapat menggambarkan dimensi pendidikan yang terdapat dalam

indikator IPM. Tabel 4.4 dapat memberikan informasi bahwa variabel APM SMP memiliki rata-rata sebesar 77,26 dengan standar deviasi 5,73. Nilai minimum dan maksimum dari variabel APM SMP masing-masing sebesar 68,51 persen yang terdapat di Kabupaten Binjai dan 91,42 persen yang terdapat di Kabupaten Dairi. Kabupaten Binjai merupakan wilayah yang memiliki APM SMP terendah dikarenakan pemerintah belum mampu menyediakan sarana dan prasarana sekolah yang memadai. Kabupaten Dairi merupakan salah satu wilayah yang memiliki tingkat pendidikan yang cukup tinggi khususnya sekolah menengah pertama (SMP), sehingga hal ini mencerminkan bahwa penduduk Kabupaten Dairi memiliki kesadaran yang tinggi terhadap pendidikan.

Variabel x_3 adalah Angka Partisipasi Murni (APM) SMA. APM SMA merupakan salah satu faktor yang diduga memengaruhi IPM dalam penelitian ini. APM SMA menunjukkan proporsi anak sekolah pada satu kelompok usia 16-18 tahun yang bersekolah pada jenjang yang sesuai dengan kelompok usianya. Semakin tinggi APM SMA menunjukkan bahwa sebagian besar penduduk di suatu wilayah mampu bersekolah tepat waktu pada jenjang SMA, sehingga hal ini akan berdampak positif pada dimensi pendidikan yang merupakan salah satu indikator dari IPM. APM SMA memiliki rata-rata sebesar 67,21 dengan standar deviasi sebesar 8,84. Nilai minimum dari APM SMA diperoleh Kabupaten Nias yaitu sebesar 47,55 persen dan nilai maksimum dari APM SMA terdapat pada Kota Padangsidimpuan dengan perolehan sebesar 87,53 persen. Kabupaten Nias mendapatkan APM SMA terendah disebabkan oleh sarana fisik pendidikan, tenaga guru PNS yang masih belum memadai. Hal ini didukung oleh informasi dari Kepala Dinas Pendidikan Kabupaten Nias bahwa jumlah guru PNS tingkat SMA hanya 64 orang dan masih kekurangan 133 orang.

Variabel x_4 adalah Tingkat Partisipasi Angkatan Kerja (TPAK) yang diduga memengaruhi IPM. Semakin tinggi TPAK maka semakin tinggi pula pasokan tenaga kerja yang tersedia untuk memproduksi barang dan jasa dalam suatu kegiatan

perekonomian. Hal ini dapat berdampak positif pada pertumbuhan ekonomi, dimana indikator pertumbuhan ekonomi merupakan salah satu faktor yang berpengaruh terhadap IPM sehingga dengan semakin baiknya pertumbuhan ekonomi maka IPM yang diperoleh akan semakin tinggi. Tabel 4.4 dapat memberikan informasi bahwa variabel TPAK memiliki rata-rata sebesar 72,80 dengan standar deviasi 9,66. Nilai minimum dan maksimum dari variabel TPAK masing-masing sebesar 57,72 persen terdapat di Kabupaten Batu Bara dan 88,38 persen terdapat di Kabupaten Samosir. Salah satu faktor yang memengaruhi tinggi rendahnya nilai TPAK adalah pendidikan. Semakin tinggi pendidikan terakhir yang ditamatkan, maka akan semakin tinggi pula nilai TPAK. Kabupaten Batu Bara menjadi wilayah dengan nilai TPAK terendah karena terdapat kecenderungan bahwa sebesar 45,28 persen penduduk usia 15 tahun ke atas yang termasuk angkatan kerja berpendidikan tertinggi yang ditamatkan yaitu tidak/belum pernah sekolah dasar (SD), tidak/belum taman SD/tamat SD.

Kabupaten Samosir memiliki nilai TPAK tertinggi di Sumatera Utara pada tahun 2015 yaitu sebesar 88,38 persen. Walaupun nilai TPAK di Kabupaten Samosir tertinggi dibandingkan dengan wilayah lainnya, nilai TPAK di wilayah tersebut mengalami penurunan dari tahun sebelumnya (tahun 2014). Hal ini disebabkan oleh kecenderungan perpindahan penduduk terutama penduduk usia kerja keluar daerah untuk mencari pekerjaan yang lebih baik atau mungkin tingkat keahlian maupun pekerjaan yang diminati tidak tersedia di Kabupaten Samosir. Penurunan nilai TPAK di Kabupaten Samosir juga dapat disebabkan oleh sebagian penduduk setempat ingin melanjutkan sekolah yang lebih tinggi dan keterbatasan lapangan kerja beberapa ibu rumah tangga yang dulunya bekerja membantu suami, memilih menjadi ibu rumah tangga saja untuk mengurus anak-anaknya.

Variabel x_5 merupakan Tingkat Pengangguran Terbuka (TPT). TPT menunjukkan jumlah pengangguran yang terdapat di suatu wilayah. Menurut Sadono Sukirno, efek dari semakin tinggi

tingkat pengangguran menunjukkan bahwa sebagian besar penduduk di suatu wilayah tidak mampu memenuhi kebutuhan hidup sehari-hari. Hal ini akan berdampak buruk pada dimensi standar hidup layak yang merupakan salah satu indikator dari IPM atau dapat dikatakan bahwa apabila dimensi standar hidup layak buruk maka IPM yang diperoleh juga akan semakin rendah. TPT memiliki rata-rata sebesar 5,84 persen dengan standar deviasi sebesar 3,31. Nilai minimum dari TPT diperoleh Kabupaten Nias Selatan yaitu sebesar 0,40 persen dan nilai maksimum dari TPT terdapat pada Kabupaten Labuhan Batu dengan perolehan sebesar 11,39 persen. Hal ini membuktikan bahwa pada Kabupaten Labuhan Batu masih banyak penduduk yang menganggur. Kabupaten Nias Selatan memiliki nilai TPT terendah karena banyaknya pekerja keluarga yang biasanya terdiri dari istri/suami dan anak-anaknya yang membantu usaha rumah tangga, dimana mayoritas usaha rumah tangga di Kabupaten Nias Selatan adalah sektor pertanian.

Penyebab Kabupaten Labuhan Batu memiliki nilai TPT tertinggi adalah adanya penurunan penyerapan tenaga kerja. Terdapat tiga sektor lapangan pekerjaan utama yaitu sektor primer, sekunder dan tersier. Sektor primer terdiri dari sektor pertanian dan pertambangan/penggalian. Sektor sekunder terdiri dari sektor industri pengolahan, sektor listrik, gas dan air bersih serta sektor konstruksi/bangunan, sedangkan sektor tersier terdiri dari sektor perdagangan, angkutan, keuangan dan jasa-jasa lainnya. Berdasarkan data yang diperoleh, pada periode 2012 sektor primer menyerap tenaga kerja sekitar 50,08 persen dan tiga tahun kemudian (tahun 2015) mengalami penurunan secara fluktuatif menjadi sekitar 49,55 persen dari total penduduk yang bekerja. Sektor sekunder juga mengalami perubahan dari 7,04 persen pada tahun 2012 meningkat menjadi 9,65 persen pada tahun 2015. Proses peningkatannya masih sedikit dan cenderung stagnan pada periode 2012 hingga 2015. Sementara itu, sektor tersier mengalami penurunan yang sama yaitu dari sekitar 42,90 persen pada tahun 2012 menjadi 40,80 persen pada tahun 2015.

Variabel x_6 merupakan Angka Harapan Hidup (AHH), dimana salah satu faktor yang memengaruhi IPM. Berdasarkan Tabel 4.4 diketahui bahwa variabel AHH memiliki rata-rata sebesar 68,07 dengan standar deviasi 2,56. Rata-rata AHH di Sumatera Utara masih rendah dibandingkan dengan AHH di Indonesia yaitu sebesar 70,78. Selain itu juga AHH Sumatera Utara tergolong rendah apabila dibandingkan dengan AHH wilayah Pulau Sumatera yang lainnya. Nilai minimum dan maksimum dari variabel AHH masing-masing sebesar 61,58 terdapat di Kabupaten Mandailing Natal dan 72,29 terdapat di Kota Pematangsiantar. Salah satu penyebab Kabupaten Mandailing Natal memiliki AHH terendah adalah lebih dari 50 persen rumah tangga menggunakan tempat pembuangan tinja di kolam/sawah/sungai/danau/laut. Hal ini nantinya menyebabkan air menjadi tercemar. Apabila air sudah tercemar, maka akan menimbulkan berbagai penyakit menular. Dengan demikian dapat dikatakan bahwa penduduk di Kabupaten Mandailing Natal masih belum menerapkan gaya hidup yang bersih, sehingga diharapkan pemerintah memberikan sosialisasi untuk menerapkan gaya hidup bersih dan sehat kepada penduduk khususnya Kabupaten Mandailing Natal.

Kota Pematangsiantar memiliki AHH tertinggi di Sumatera Utara karena sebagian besar penduduk telah menerapkan gaya hidup bersih dan sehat. Hal ini dapat dibuktikan dari data yang telah diperoleh, tempat pembuangan akhir tinja sebagian besar rumah tangga di Kota Pematangsiantar adalah tangki/SPAL yaitu sebesar 91,51 persen. Selain itu, sebesar 69,56 persen penduduk sudah menggunakan air leding meteran sebagai sumber air yang layak diminum.

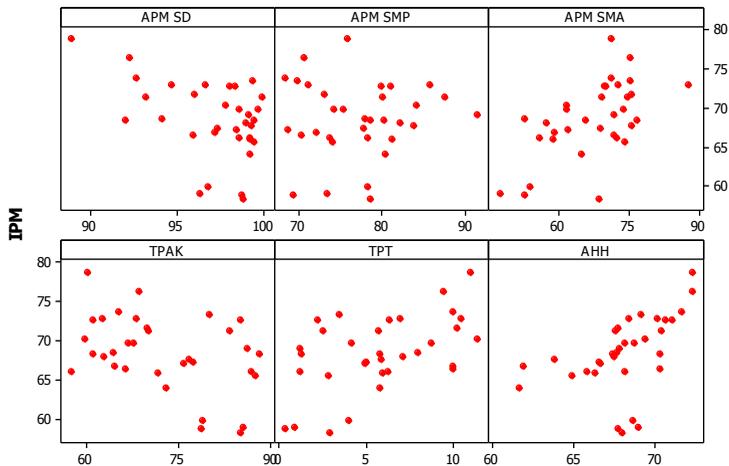
4.2 Pemodelan Indeks Pembangunan Manusia

Sebelum melakukan pemodelan Indeks Pembangunan Manusia (IPM) dengan faktor-faktor yang memengaruhinya, terdapat beberapa langkah yang perlu dilakukan antara lain melihat hubungan antara IPM dengan masing-masing variabel prediktor, memilih titik knot optimum berdasarkan nilai GCV paling minimum, melakukan pengujian signifikansi parameter model baik

uji serentak maupun uji parsial, melakukan pengujian asumsi residual.

4.2.1 Analisis Faktor-Faktor yang Memengaruhi Indeks Pembangunan Manusia

Sebelum melakukan pemodelan antara IPM dengan variabel prediktor yang diduga memengaruhinya, maka perlu terlebih dahulu melihat hubungan antara IPM dengan masing-masing variabel prediktor. Gambar 4.1 menunjukkan hubungan antara IPM Sumatera Utara dengan 6 variabel prediktor yang memengaruhinya.



Gambar 4.1 Scatterplot antara IPM dengan 6 Variabel Prediktor

Apabila dilihat dari Gambar 4.1, secara visual hubungan antara IPM Sumatera Utara dengan enam variabel prediktor tidak membentuk pola tertentu, sehingga keenam variabel prediktor merupakan komponen nonparametrik dan dalam melakukan pemodelan digunakan metode regresi nonparametrik Spline. Model Spline yang digunakan adalah Spline Truncated Linear dengan satu titik knot, dua titik knot, tiga titik knot dan kombinasi titik knot.

4.2.2 Pemilihan Titik Knot Optimum

Titik knot merupakan titik perpaduan dimana terjadi perubahan pola data. Model regresi nonparametrik Spline terbaik diperoleh dari pemilihan titik knot optimum. Pemilihan titik knot optimum menggunakan metode *Generalized Cross Validation* (GCV). Semakin nilai GCV maka semakin optimum titik knot yang dipilih

1. Pemilihan Titik Knot Optimum dengan Satu Titik Knot

Pemilihan titik knot optimum diawali dengan menggunakan satu titik knot, sehingga dengan menggunakan satu titik knot pada variabel-variabel yang memengaruhi IPM Sumatera Utara diharapkan dapat menemukan nilai GCV yang paling minimum dan nantinya akan menghasilkan model Spline terbaik. Berikut ini merupakan estimasi model regresi nonparametrik Spline dengan satu titik knot pada kasus IPM Sumatera Utara.

$$y = \gamma_0 + \gamma_1 x_1 + \gamma_2 (x_1 - K_1)_+^1 + \gamma_3 x_2 + \gamma_4 (x_2 - K_2)_+^1 + \gamma_5 x_3 + \gamma_6 (x_3 - K_3)_+^1 + \gamma_7 x_4 + \gamma_8 (x_4 - K_4)_+^1 + \gamma_9 x_5 + \gamma_{10} (x_5 - K_5)_+^1 + \gamma_{11} x_6 + \gamma_{12} (x_6 - K_6)_+^1$$

Tabel 4.5 menyajikan sepuluh nilai GCV di sekitar nilai GCV paling minimum untuk model regresi nonparametrik Spline. Tabel 4.5 menunjukkan bahwa nilai GCV minimum dengan satu titik knot sebesar 13,92. Berdasarkan nilai GCV paling minimum pada variabel x_1 yang merupakan APM SD adalah pada nilai 96,17. Hal ini berarti bahwa pada nilai tersebut terjadi perubahan perilaku pada fungsi yang dihasilkan. Variabel x_2 merupakan APM SMP, titik knot diperoleh pada saat 83,47 yang berarti bahwa fungsi memiliki perubahan pada saat x_2 mencapai nilai 83,47. Variabel x_3 menunjukkan APM SMA yang mendapatkan titik knot pada angka 73,65. Hal ini berarti bahwa pada saat APM SMA bernilai 73,65 maka fungsi mengalami perubahan perilaku. Selanjutnya titik knot pada variabel x_4 menunjukkan TPAK terjadi pada saat 77,74 yang dapat diartikan bahwa pada nilai tersebut terjadi perubahan perilaku pada fungsi yang dihasilkan.

Variabel x_5 dan x_6 masing-masing menunjukkan TPT dan AHH. Titik knot pada variabel x_5 terjadi pada saat 7,57, sedangkan titik knot pada variabel x_6 terjadi pada saat 68,57. Hal ini berarti bahwa fungsi mengalami perubahan perilaku pada saat TPT bernilai 7,57. Selain itu, hal serupa juga terjadi pada saat AHH bernilai 68,57 maka fungsi mengalami perubahan perilaku.

Tabel 4.5 Nilai GCV dengan Satu Titik Knot

No	Knot						GCV
	x_1	x_2	x_3	x_4	x_5	x_6	
1	95,27	81,60	70,39	75,24	6,68	67,70	15,00
2	95,49	82,07	71,21	75,86	6,90	67,91	14,52
3	95,72	82,54	72,02	76,49	7,12	68,13	14,19
4	95,94	83,00	72,84	77,11	7,35	68,35	13,96
5	96,17	83,47	73,65	77,74	7,57	68,57	13,92
6	96,39	83,94	74,47	78,36	7,80	68,79	14,05
7	96,62	84,40	75,29	78,99	8,02	69,01	14,18
8	96,84	84,87	76,10	79,62	8,25	69,23	14,31
9	97,07	85,34	76,92	80,24	8,47	69,44	14,47
10	97,29	85,81	77,73	80,87	8,69	69,66	14,65

Nilai GCV minimum untuk model regresi nonparametrik Spline linier dengan satu titik knot akan dibandingkan dengan model regresi nonparametrik Spline linier dengan dua knot, tiga knot dan kombinasi knot untuk mendapatkan model terbaik.

2. Pemilihan Titik Knot Optimum dengan Dua Titik Knot

Setelah dilakukan pemilihan titik knot dengan menggunakan satu titik knot, langkah selanjutnya adalah melakukan pemilihan titik knot dengan menggunakan dua titik knot. Berikut ini merupakan estimasi model regresi nonparametrik Spline dengan dua titik knot pada kasus IPM Sumatera Utara.

$$\begin{aligned}
 y = & \gamma_0 + \gamma_1 x_1 + \gamma_2 (x_1 - K_1)_+^1 + \gamma_3 (x_1 - K_2)_+^1 + \gamma_4 x_2 + \gamma_5 (x_2 - K_3)_+^1 + \gamma_6 (x_2 - K_4)_+^1 + \\
 & \gamma_7 x_3 + \gamma_8 (x_3 - K_5)_+^1 + \gamma_9 (x_3 - K_6)_+^1 + \gamma_{10} x_4 + \gamma_{11} (x_4 - K_7)_+^1 + \gamma_{12} (x_4 - K_8)_+^1 + \\
 & \gamma_{13} x_5 + \gamma_{14} (x_5 - K_9)_+^1 + \gamma_{15} (x_5 - K_{10})_+^1 + \gamma_{16} x_6 + \gamma_{17} (x_6 - K_{11})_+^1 + \gamma_{18} (x_6 - K_{12})_+^1
 \end{aligned}$$

Model regresi nonparametrik Spline linier dengan dua titik knot yang dibentuk seperti persamaan diatas maka diperlukan dua titik knot yang optimum pada setiap variabel prediktor. Tabel 4.6 merupakan sepuluh nilai GCV minimum untuk model regresi nonparametrik Spline linier menggunakan dua titik knot.

Tabel 4.6 Nilai GCV dengan Dua Titik Knot

No	Knot						GCV
	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	
1	88,97	68,52	47,55	57,72	0,40	61,58	14,83
	98,87	89,08	83,45	85,25	10,26	71,19	
2	88,97	68,52	47,55	57,72	0,40	61,58	14,90
	99,09	89,55	84,26	85,87	10,49	71,41	
3	88,97	68,52	47,55	57,72	0,40	61,58	15,13
	99,32	90,01	85,08	86,5	10,71	71,63	
4	88,97	68,52	47,55	57,72	0,40	61,58	15,23
	99,54	90,48	85,89	87,12	10,94	71,85	
5	88,97	68,52	47,55	57,72	0,40	61,58	14,82
	99,77	90,95	86,71	87,75	11,16	72,07	
6	88,97	68,52	47,55	57,72	0,40	61,58	10,86
	100	91,42	87,53	88,38	11,39	72,29	
7	89,19	68,98	48,36	58,34	0,62	61,79	18,95
	89,42	69,45	49,18	58,97	0,84	62,01	
8	89,19	68,98	48,36	58,34	0,62	61,79	18,95
	89,64	69,92	49,99	59,59	1,07	62,23	
9	89,19	68,98	48,36	58,34	0,62	61,79	22,61
	89,87	70,38	50,81	60,22	1,29	62,45	
10	89,19	68,98	48,36	58,34	0,62	61,79	23,39
	90,09	70,85	51,62	60,84	1,52	62,67	

Berdasarkan Tabel 4.6, diketahui bahwa nilai GCV minimum yang diperoleh dari pemodelan regresi nonparametrik Spline linier dengan dua titik knot adalah 10,86 dengan titik-titik knot optimum sebagai berikut.

$K_1=88,97$; $K_2=100$

$K_3=68,52$; $K_4=91,42$

$K_5=47,55$; $K_6=87,53$

$$\begin{aligned}
K_7 &= 57,72 ; K_8 = 88,38 \\
K_9 &= 0,40 ; K_{10} = 11,39 \\
K_{11} &= 61,58 ; K_{12} = 72,29
\end{aligned}$$

Nilai GCV pada pemilihan titik knot dengan dua titik knot lebih kecil daripada nilai GCV pada pemilihan titik knot dengan satu titik knot. Hal ini mencerminkan bahwa model regresi non-parametrik Spline linier dengan dua titik knot lebih baik dibandingkan model regresi nonparametrik Spline linier dengan satu titik knot.

3. Pemilihan Titik Knot Optimum dengan Tiga Titik Knot

Berikut ini merupakan estimasi model regresi nonparametrik Spline dengan tiga titik knot pada kasus IPM Sumatera Utara.

$$\begin{aligned}
y &= \gamma_0 + \gamma_1 x_1 + \gamma_2 (x_1 - K_1)_+^1 + \gamma_3 (x_1 - K_2)_+^1 + \gamma_4 (x_1 - K_3)_+^1 + \gamma_5 x_2 + \gamma_6 (x_2 - K_4)_+^1 + \\
&\quad \gamma_7 (x_2 - K_5)_+^1 + \gamma_8 (x_2 - K_6)_+^1 + \gamma_9 x_3 + \gamma_{10} (x_3 - K_7)_+^1 + \gamma_{11} (x_3 - K_8)_+^1 + \\
&\quad \gamma_{12} (x_3 - K_9)_+^1 + \gamma_{13} x_4 + \gamma_{14} (x_4 - K_{10})_+^1 + \gamma_{15} (x_4 - K_{11})_+^1 + \gamma_{16} (x_4 - K_{12})_+^1 + \\
&\quad \gamma_{14} x_5 + \gamma_{14} (x_5 - K_{13})_+^1 + \gamma_{15} (x_5 - K_{14})_+^1 + \widehat{\gamma_{16}} (x_5 - K_{15})_+^1 + \widehat{\gamma_{17}} x_6 + \\
&\quad \widehat{\gamma_{18}} (x_6 - K_{16})_+^1 + \widehat{\gamma_{19}} (x_6 - K_{17})_+^1 + \widehat{\gamma_{20}} (x_6 - K_{18})_+^1
\end{aligned}$$

Model regresi nonparametrik Spline linier dengan tiga titik knot yang dibentuk seperti persamaan diatas maka diperlukan tiga titik knot yang optimum pada setiap variabel prediktor. Tabel 4.7 merupakan sepuluh nilai GCV minimum untuk model regresi nonparametrik Spline linier menggunakan tiga titik knot. Berdasarkan Tabel 4.7 diketahui bahwa nilai GCV minimum yang diperoleh dari pemodelan regresi nonparametrik Spline linier dengan tiga titik knot adalah 12,85 dengan titik-titik knot optimum sebagai berikut.

$$\begin{aligned}
K_1 &= 97,97 ; K_2 = 98,42 ; K_3 = 98,64 \\
K_4 &= 88,21 ; K_5 = 88,14 ; K_6 = 88,61 \\
K_7 &= 80,18 ; K_8 = 81,81 ; K_9 = 82,63 \\
K_{10} &= 82,74 ; K_{11} = 84,00 ; K_{12} = 84,62 \\
K_{13} &= 9,37 ; K_{14} = 9,82 ; K_{15} = 10,04 \\
K_{16} &= 70,32 ; K_{17} = 76,76 ; K_{18} = 70,97
\end{aligned}$$

Tabel 4.7 Nilai GCV dengan Tiga Titik Knot

No	Knot						GCV
	x_1	x_2	x_3	x_4	x_5	x_6	
1	97,97	87,21	80,18	82,74	9,37	70,32	15,81
	98,19	87,68	81,00	83,37	9,59	70,54	
	98,87	89,08	83,45	85,25	10,26	71,19	
2	97,97	87,21	80,18	82,74	9,37	70,32	17,96
	98,19	87,68	81,00	83,37	9,59	70,54	
	99,09	89,55	84,26	85,87	10,49	71,41	
3	97,97	87,21	80,18	82,74	9,37	70,32	18,05
	98,19	87,68	81,00	83,37	9,59	70,54	
	99,32	90,01	85,08	86,5	10,71	71,63	
4	97,97	87,21	80,18	82,74	9,37	70,32	22,68
	98,19	87,68	81,00	83,37	9,59	70,54	
	99,54	90,48	85,89	87,12	10,94	71,85	
5	97,97	87,21	80,18	82,74	9,37	70,32	17,34
	98,19	87,68	81,00	83,37	9,59	70,54	
	99,77	90,95	86,71	87,75	11,16	72,07	
6	97,97	87,21	80,18	82,74	9,37	70,32	12,85
	98,42	88,14	81,81	84,00	9,82	70,76	
	98,64	88,61	82,63	84,62	10,04	70,97	
7	97,97	87,21	80,18	82,74	9,37	70,32	18,84
	98,42	88,14	81,81	84,00	9,82	70,76	
	98,87	89,08	83,45	85,25	10,26	71,19	
8	97,97	87,21	80,18	82,74	9,37	70,32	25,79
	98,42	88,14	81,81	84,00	9,82	70,76	
	99,09	89,55	84,26	85,87	10,49	71,41	
9	97,97	87,21	80,18	82,74	9,37	70,32	28,13
	98,42	88,14	81,81	84,00	9,82	70,76	
	99,32	90,01	85,08	86,5	10,71	71,63	
10	97,97	87,21	80,18	82,74	9,37	70,32	27,42
	98,42	88,14	81,81	84,00	9,82	70,76	
	99,54	90,48	85,89	87,12	10,94	71,85	

4. Pemilihan Titik Knot Optimum dengan Kombinasi Titik Knot

Kombinasi knot adalah kombinasi antara satu titik knot, dua titik knot dan tiga titik knot. Kombinasi ini dilakukan karena masing-masing variabel prediktor mengalami perubahan pola perilaku data yang berbeda-beda terhadap variabel respon. Selain itu, kombinasi titik knot dilakukan juga untuk memilih titik knot optimum dengan melihat nilai GCV minimum. Hasil GCV kombinasi knot akan dibandingkan dengan hasil GCV model regresi nonparametrik Spline sebelumnya. Berikut ini merupakan sepuluh nilai diantara nilai GCV minimum yang disajikan dalam Tabel 4.8. Berdasarkan Tabel 4.8, diketahui bahwa nilai GCV minimum dengan menggunakan kombinasi knot (2,1,2,3,3,3) diperoleh sebesar 7,94. Titik knot yang dihasilkan pada variabel APM SD (x_1) adalah 88,97 dan 100. Hal ini menunjukkan bahwa fungsi yang dihasilkan akan mengalami perubahan perilaku pada nilai tersebut. Variabel x_2 merupakan APM SMP memperoleh titik knot pada saat 83,47. Hal ini berarti bahwa fungsi memiliki perubahan perilaku pada saat x_2 mencapai nilai 83,47. Variabel x_3 adalah variabel APM SMA yang memiliki titik knot 47,55 dan 87,53. Nilai tersebut dapat diartikan bahwa pada saat APM SMA bernilai 47,55 dan 87,53 maka fungsi mengalami perubahan perilaku.

Variabel TPAK (x_4) memperoleh titik knot sebesar 82,74, 84 dan 84,62. Nilai tersebut dapat diartikan bahwa apabila TPAK bernilai 82,74, 84 dan 84,62 maka fungsi akan mengalami perubahan perilaku. Titik knot yang diperoleh pada variabel TPT (x_5) yaitu 9,37, 9,82 dan 10,04. Hal ini menunjukkan apabila TPT bernilai 9,37, 9,82 dan 10,04 maka fungsi mengalami perubahan perilaku. Variabel x_6 menunjukkan variabel AHH dengan titik knot yang diperoleh sebesar 70,32, 70,76 dan 70,97. Nilai tersebut berarti bahwa pada saat AHH bernilai 70,32, 70,76 dan 70,97 maka terjadi perubahan perilaku pada fungsi yang dihasilkan.

Tabel 4.8 Nilai GCV dengan Kombinasi Titik Knot

No	Knot						GCV
	x ₁	x ₂	x ₃	x ₄	x ₅	x ₆	
1				82,74	9,37		
	88,97		47,55	84	9,82		
	100	83,47	87,53	84,62	10,04	68,57	11,58
2				82,74	9,37		
	88,97		47,55	84	9,82	61,58	
	100	83,47	87,53	84,62	10,04	72,29	15,62
3				82,74	9,37	70,32	
	88,97		47,55	84	9,82	70,76	
	100	83,47	87,53	84,62	10,04	70,97	7,94
4			80,18				
	88,97		81,81				
	100	83,47	82,63	77,74	7,57	68,57	14,52
5			80,18				
	88,97		81,81			61,58	
	100	83,47	82,63	77,74	7,57	72,29	13,91
6			80,18			70,32	
	88,97		81,81			70,76	
	100	83,47	82,63	77,74	7,57	70,97	14,00
7			80,18				
	88,97		81,81		0,40		
	100	83,47	82,63	77,74	11,39	68,57	13,34
8			80,18				
	88,97		81,81		0,40	61,58	
	100	83,47	82,63	77,74	11,39	72,29	12,83
9			80,18			70,32	
	88,97		81,81		0,40	70,76	
	100	83,47	82,63	77,74	11,39	70,97	12,83
10			80,18		9,37		
	88,97		81,81		9,82		
	100	83,47	82,63	77,74	10,04	68,57	13,11

Setelah diperoleh nilai GCV minimum dalam pemilihan titik knot dengan menggunakan satu titik knot, dua titik knot, tiga titik knot dan kombinasi titik knot, maka langkah selanjutnya yang dilakukan adalah memilih model terbaik berdasarkan nilai GCV paling minimum. Perbandingan nilai GCV model regresi nonpa-

rametrik Spline dengan satu titik knot, dua titik knot, tiga titik knot dan kombinasi knot disajikan dalam Tabel 4.9.

Tabel 4.9 Perbandingan Nilai GCV Berbagai Knot

Model	GCV
1 Knot	13,92
2 Knot	10,86
3 Knot	12,85
Kombinasi Knot (2,1,2,3,3,3)	7,94

Berdasarkan Tabel 4.9, nilai GCV paling minimum diberikan oleh model regresi nonparametrik Spline dengan menggunakan kombinasi knot yaitu sebesar 7,94. Oleh karena itu, model regresi nonparametrik Spline yang digunakan untuk memodelkan Indeks Pembangunan Manusia di Sumatera Utara dengan seluruh komponen nonparametrik adalah dengan kombinasi dua titik knot untuk variabel x_1 , satu titik knot untuk variabel x_2 , dua titik knot untuk variabel x_3 dan tiga titik knot masing-masing untuk variabel x_4 , x_5 , x_6 .

Pemodelan regresi nonparametrik Spline menggunakan titik knot optimal yang diperoleh dari nilai GCV paling minimum. Model regresi nonparametrik Spline terbaik dapat dituliskan sebagai berikut.

$$\begin{aligned}
 y = & 0,009 + 0,485x_1 - 0,396(x_1 - 88,97)_+^1 - 7,358 \times 10^{-10}(x_1 - 100)_+^1 - 0,279x_2 + \\
 & 1,215(x_2 - 83,47)_+^1 + 0,342x_3 - 0,129(x_3 - 47,55)_+^1 + 1,787 \times 10^{-10}(x_3 - 87,53)_+^1 - \\
 & 0,088x_4 + 8,073(x_4 - 82,74)_+^1 - 32,627(x_4 - 84)_+^1 + 27,009(x_4 - 84,62)_+^1 + \\
 & 0,958x_5 - 65,261(x_5 - 9,37)_+^1 + 179,589(x_5 - 9,82)_+^1 - 113,817(x_5 - 10,04)_+^1 + \\
 & 0,382x_6 + 41,673(x_6 - 70,32)_+^1 - 11,852(x_6 - 70,76)_+^1 + 72,817(x_6 - 70,97)_+^1
 \end{aligned}$$

4.2.3 Pengujian Signifikansi Parameter Model Regresi Nonparametrik Spline

Pengujian signifikansi parameter dilakukan untuk mengetahui apakah parameter yang telah didapatkan dari hasil pemodelan dengan regresi nonparametrik Spline berpengaruh secara signifikan terhadap variabel IPM atau tidak. Pengujian ini dilakukan dengan dua tahap yaitu uji secara serentak dan uji secara parsial. Apabila pada uji secara serentak, parameter menunjukkan

pengaruh yang signifikan terhadap IPM maka akan dilanjutkan pada uji secara parsial. Uji secara parsial dilakukan untuk mengetahui pengaruh signifikansi dari tiap-tiap parameter terhadap IPM.

1. Pengujian Secara Serentak

Pengujian secara serentak dilakukan pada parameter model regresi nonparametrik Spline terhadap IPM secara serentak. Berikut ini merupakan hipotesis pengujian parameter secara serentak.

$$H_0 : \gamma_1 = \gamma_2 = \dots = \gamma_{20} = 0$$

$$H_1 : \text{minimal ada satu } \gamma_j \neq 0, j = 1, 2, \dots, 20$$

Statistik uji menggunakan uji F dan hasil pengujian parameter secara serentak berupa tabel ANOVA diberikan oleh Tabel 4.10.

Tabel 4.10 Tabel ANOVA Pengujian Secara Serentak

Sumber Variasi	df	Sum of Square	Mean Square	F_{hitung}	P-value
Regresi	20	691,08	34,55	6,72	0,0007
Error	12	61,66	5,13		
Total	32	752,74			

Berdasarkan ANOVA pada Tabel 4.10, dapat diperoleh nilai F_{hitung} dan $P\text{-value}$ masing-masing sebesar 6,72 dan 0,0007. Keputusan yang diambil dari pengujian parameter secara serentak dengan menggunakan taraf signifikansi (α) sebesar 0,05 adalah tolak H_0 karena apabila nilai F_{hitung} dibandingkan dengan F_{tabel} (0,05;20;12) sebesar 2,28, maka F_{hitung} lebih besar daripada F_{tabel} . Selain itu, apabila $P\text{-value}$ dibandingkan dengan $\alpha=0,05$ diperoleh hasil bahwa $P\text{-value}$ kurang dari α . Hal ini berarti bahwa minimal ada satu parameter pada model yang berpengaruh signifikan terhadap variabel respon.

2. Pengujian Secara Parsial

Langkah selanjutnya untuk mengetahui parameter mana saja yang berpengaruh signifikan terhadap variabel respon atau dalam hal ini adalah IPM, maka perlu dilanjutkan ke tahap pengujian secara parsial.

Hipotesis untuk pengujian parameter model secara parsial adalah sebagai berikut.

$$H_0 : \gamma_j = 0$$

$$H_1 : \gamma_j \neq 0, j = 1, 2, \dots, 20$$

Statistik uji menggunakan uji t dan hasil dari pengujian parameter model regresi secara parsial disajikan dalam Tabel 4.11 sebagai berikut.

Tabel 4.11 Hasil Pengujian Secara Parsial

Variabel	Parameter	Estimasi	P-value	Kesimpulan
x_1	γ_1	0,485	0,104	Tidak Signifikan
	γ_2	-0,396	0,493	Tidak Signifikan
	γ_3	$-7,358 \times 10^{-10}$	0,016	Signifikan
x_2	γ_4	-0,279	0,111	Tidak Signifikan
	γ_5	1,215	0,017	Signifikan
x_3	γ_6	0,342	0,170	Tidak Signifikan
	γ_7	-0,129	0,542	Tidak Signifikan
	γ_8	$1,787 \times 10^{-10}$	0,006	Signifikan
x_4	γ_9	-0,088	0,446	Tidak Signifikan
	γ_{10}	8,073	0,043	Signifikan
	γ_{11}	-32,627	0,007	Signifikan
	γ_{12}	27,009	0,004	Signifikan
x_5	γ_{13}	0,958	0,035	Signifikan
	γ_{14}	-65,261	0,027	Signifikan
	γ_{15}	179,589	0,048	Signifikan
	γ_{16}	-113,817	0,074	Tidak Signifikan
x_6	γ_{17}	0,382	0,259	Tidak Signifikan
	γ_{18}	41,673	0,001	Signifikan
	γ_{19}	-111,852	0,004	Signifikan
	γ_{20}	72,817	0,008	Signifikan

Berdasarkan Tabel 4.11, diketahui bahwa dengan taraf signifikansi (α) sebesar 0,05 dari 20 parameter menunjukkan bahwa hanya terdapat 8 parameter yang tidak signifikan diantaranya $\gamma_1, \gamma_2, \gamma_4, \gamma_6, \gamma_7, \gamma_9, \gamma_{16}, \gamma_{17}$ karena *P-value* yang dihasilkan lebih dari α . Meskipun terdapat 8 parameter yang tidak signifikan, keenam variabel disimpulkan berpengaruh terhadap IPM.

4.2.4 Pengujian Asumsi Residual

Pengujian asumsi residual dilakukan untuk mengetahui kelayakan suatu model regresi apakah residual yang telah terbentuk memenuhi asumsi identik, independen dan berdistribusi Normal (IIDN) atau tidak. Apabila suatu model regresi dengan kriteria model terbaik dan parameter signifikan, namun residual tidak memenuhi asumsi IIDN maka model tersebut tidak layak digunakan untuk memodelkan variabel respon atau dengan kata lain model regresi nonparametrik Spline tidak layak menggambarkan hubungan antara variabel prediktor dengan variabel respon.

1. Uji Identik

Pengujian asumsi residual identik terpenuhi apabila tidak terjadi kasus heteroskedastisitas atau varians residual dari model harus homogen. Pemeriksaan asumsi residual identik dapat dilakukan dengan menggunakan uji *Glejser*. Berikut ini merupakan hipotesis yang digunakan dalam uji *Glejser*.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{33}^2 = \sigma^2$$

$$H_1 : \text{minimal terdapat satu } \sigma_i^2 \neq \sigma^2, i = 1, 2, \dots, 33$$

Hasil uji *Glejser* disajikan dalam Tabel 4.12 sebagai berikut.

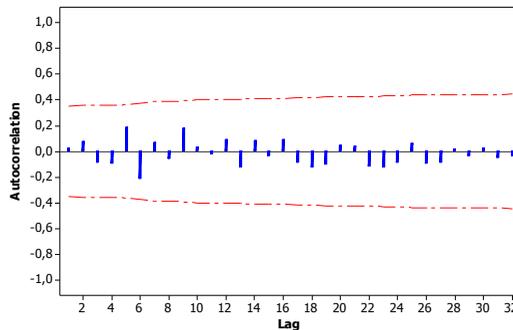
Tabel 4.12 ANOVA Uji *Glejser*

Sumber Variasi	Df	Sum of Square	Mean Square	F_{hitung}	P-value
Regresi	20	15,74	0,78	0,81	0,66
Error	12	11,53	0,96		
Total	32	27,27			

Berdasarkan Tabel 4.12, diketahui bahwa *P-value* sebesar 0,66 dengan taraf signifikansi (α) sebesar 0,05 maka diperoleh keputusan gagal tolak H_0 karena *P-value* lebih besar dibandingkan α . Selain itu, nilai $F_{hitung} < F_{tabel}$ (0,05;20;12) yaitu $0,81 < 2,28$ sehingga dapat disimpulkan bahwa tidak terjadi kasus heteroskedastisitas atau varians antar residual homogen. Hal ini berarti bahwa asumsi residual identik telah terpenuhi.

2. Uji Independen

Asumsi residual selanjutnya yang harus terpenuhi adalah independen. Residual yang independen apabila tidak terjadi autokorelasi antar residual. Pemeriksaan asumsi residual independen dapat menggunakan plot *Autocorrelation Function* (ACF). Apabila ada nilai autokorelasi yang keluar dari batas signifikansi maka dapat disimpulkan bahwa terjadi korelasi antar residual. Berikut merupakan hasil plot ACF yang disajikan dalam Gambar 4.2.



Gambar 4.2 Plot ACF Residual

Gambar 4.2 menunjukkan bahwa tidak terdapat nilai autokorelasi yang keluar dari batas signifikansi, sehingga dapat disimpulkan bahwa tidak terdapat kasus autokorelasi antar residual. Hal ini berarti bahwa residual telah memenuhi asumsi independen. Selain itu, pengujian asumsi independen juga dilakukan dengan uji Durbin Watson.

Berikut merupakan hipotesis pengujian Durbin Watson.

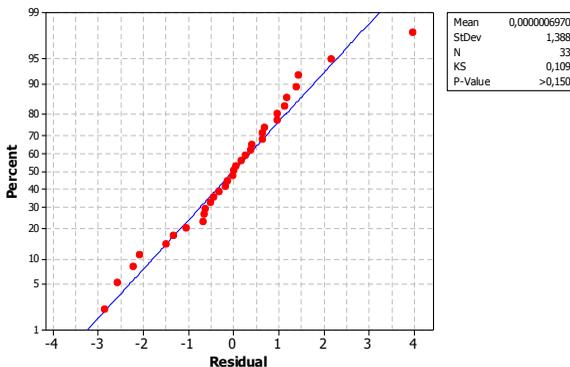
$H_0 : \rho = 0$ (Tidak terjadi autokorelasi antar residual)

$H_1 : \rho \neq 0$ (Terjadi autokorelasi antar residual)

Hasil dari pengujian Durbin Watson, diperoleh d_{hitung} sebesar 2,003. Apabila d_{hitung} dibandingkan nilai $dL=1,061$ dan $dU=1,900$ maka $d_{hitung} > dU$ atau $(4-d_{hitung}) > dU$. Dengan demikian, keputusan yang diperoleh yaitu gagal tolak H_0 yang berarti bahwa tidak ada korelasi antar residual.

3. Uji Distribusi Normal

Pengujian asumsi distribusi Normal dilakukan untuk mengetahui apakah residual telah berdistribusi Normal atau tidak. Pemeriksaan asumsi distribusi Normal dapat dilakukan dengan menggunakan uji *Kolmogorov-Smirnov*. Berikut ini adalah hasil dari uji *Kolmogorov-Smirnov* dapat ditampilkan pada Gambar 4.3.



Gambar 4.3 Normality Probability Plot Residual

Berdasarkan Gambar 4.3, secara visual residual berdistribusi Normal karena pola residual mengikuti garis normal. Keputusan uji *Kolmogorov-Smirnov* dapat diambil dengan membandingkan P-value dengan taraf signifikansi (α) sebesar 0,05.

Keputusan yang diperoleh adalah gagal tolak H_0 karena *P-value* lebih besar dari 0,150 sehingga apabila *P-value* dibandingkan dengan $\alpha=0,05$ maka *P-value* lebih besar daripada α . Hal ini berarti bahwa residual telah berdistribusi Normal. Ketiga asumsi residual telah terpenuhi, sehingga dapat disimpulkan bahwa model regresi layak digunakan untuk menggambarkan hubungan variabel prediktor dengan variabel respon.

4.3 Pembahasan

Model regresi nonparametrik Spline yang terpilih menjadi model terbaik adalah model regresi nonparametrik Spline dengan menggunakan kombinasi knot (2,1,2,3,3,3) dan diperoleh nilai R^2 sebesar 91,80 persen. Nilai tersebut berarti bahwa model terbaik yang didapatkan mampu menjelaskan variasi variabel respon yaitu IPM Sumatera Utara sebesar 91,80 persen, sedangkan sisanya koefisien determinasi 8,20 persen lainnya dijelaskan oleh variabel prediktor lain yang tidak termasuk ke dalam model. Berikut ini merupakan model akhir yang diperoleh.

$$y = 0,009 + 0,485x_1 - 0,396(x_1 - 88,97)_+^1 - 7,358 \times 10^{-10}(x_1 - 100)_+^1 - 0,279x_2 + 1,215(x_2 - 83,47)_+^1 + 0,342x_3 - 0,129(x_3 - 47,55)_+^1 + 1,787 \times 10^{-10}(x_3 - 87,53)_+^1 - 0,088x_4 + 8,073(x_4 - 82,74)_+^1 - 32,627(x_4 - 84)_+^1 + 27,009(x_4 - 84,62)_+^1 + 0,958x_5 - 65,261(x_5 - 9,37)_+^1 + 179,589(x_5 - 9,82)_+^1 - 113,817(x_5 - 10,04)_+^1 + 0,382x_6 + 41,673(x_6 - 70,32)_+^1 - 11,852(x_6 - 70,76)_+^1 + 72,817(x_6 - 70,97)_+^1$$

Model diatas telah dilakukan pengujian signifikansi parameter baik secara serentak maupun secara parsial dan diperoleh hasil bahwa semua variabel prediktor yang digunakan dalam pemodelan berpengaruh secara signifikan terhadap variabel respon. Model diatas juga telah dilakukan pengujian asumsi residual dan semua asumsi residual telah terpenuhi. Berdasarkan model diatas. Berikut ini adalah interpretasi model pada setiap variabel yang berpengaruh secara signifikan terhadap IPM Sumatera Utara.

1. Hubungan antara APM SD terhadap IPM dengan asumsi variabel prediktor lain tetap, model yang diperoleh adalah sebagai berikut.

$$y = 0,4853x_1 - 0,3969(x_1 - 88,97)_+^1 - 7,3583 \times 10^{-10} (x_1 - 100)_+^1$$

$$= \begin{cases} 0,485x_1 & ; x_1 < 88,97 \\ 0,0884x_1 + 35,3121 & ; 88,97 \leq x_1 < 100 \\ 0,0883x_1 + 35,3121 & ; x_1 \geq 100 \end{cases}$$

Interpretasi model dari hubungan antara APM SD terhadap IPM Sumatera Utara adalah apabila Kabupaten/Kota dengan APM SD bernilai lebih dari sama dengan 100 dan nilainya naik satu persen dengan asumsi variabel prediktor lain tetap, maka IPM akan naik sebesar 0,08839 persen. Wilayah yang tergolong ke dalam interval ini adalah Kabupaten Tapanuli Utara. Kabupaten Tapanuli Utara tergolong dalam wilayah dengan status IPM tinggi yaitu IPM berada pada kisaran 70 hingga 80. Tingginya nilai APM SD di suatu wilayah dapat dipengaruhi oleh faktor sosial dan ekonomi orang tua, walaupun sebagian besar penduduk Kabupaten Tapanuli Utara bermata pencaharian di sektor pertanian yang tergolong pendapatannya rendah, penduduknya masih memerhatikan pendidikan khususnya program wajib belajar 12 tahun.

2. Hubungan antara Angka Partisipasi Murni (APM) SMP terhadap IPM dengan asumsi variabel prediktor lain tetap, dimana model terbaik yang didapatkan untuk variabel APM SMP terdapat dua interval yaitu sebagai berikut.

$$y = -0,279x_2 + 1,215(x_2 - 83,47)_+^1$$

$$= \begin{cases} -0,279x_2 & ; x_2 < 83,47 \\ 0,936x_2 - 101,4577 & ; x_2 \geq 83,47 \end{cases}$$

Berdasarkan model diatas, dapat diketahui bahwa pada interval kedua yaitu pada kondisi APM SMP bernilai lebih dari sama dengan 83,47 dan naik sebesar satu persen dengan asumsi variabel prediktor lain tetap, maka IPM akan meningkat sebesar 0,936 persen. Wilayah yang termasuk ke

dalam interval kedua ini adalah Kabupaten Tapanuli Selatan, Kabupaten Labuhan Batu, Kabupaten Simalungun, Kabupaten Dairi dan Kota Padangsidempuan. Wilayah tersebut juga didominasi oleh wilayah dengan status IPM tinggi.

3. Hubungan antara APM SMA terhadap IPM dengan asumsi variabel prediktor lain tetap, maka model yang diperoleh adalah sebagai berikut.

$$y = 0,342x_3 - 0,129(x_3 - 47,55)_+^1 + 1.787 \times 10^{-10}(x_3 - 87,53)_+^1$$

$$= \begin{cases} 0,342x_3 ; x_3 < 47,55 \\ 0,213x_3 + 6,14346 ; 47,55 \leq x_3 < 87,53 \\ 0,213x_3 + 6,14345 ; x_3 \geq 87,53 \end{cases}$$

Interpretasi dari model diatas adalah pada saat APM SMA suatu Kabupaten/Kota bernilai lebih dari sama dengan 87,53 dan nilainya naik satu persen dengan asumsi variabel prediktor lain tetap, maka IPM akan naik sebesar 0,213 persen. Wilayah yang tergolong ke dalam interval ini adalah Kota Padangsidempuan. Kota Padangsidempuan merupakan wilayah yang memiliki sarana fisik pendidikan jenjang SMA yang memadai dan juga tergolong dalam wilayah dengan status IPM tinggi yaitu IPM berkisar antara 70 hingga 80.

4. Hubungan antara TPAK terhadap IPM dengan asumsi variabel prediktor lain tetap, maka model yang diperoleh adalah sebagai berikut.

$$y = -0,08x_4 + 8,07(x_4 - 82,74)_+^1 - 32,62(x_4 - 84)_+^1 + 27,009(x_4 - 84,62)_+^1$$

$$= \begin{cases} -0,08x_4 ; x_4 < 82,74 \\ 7,98x_4 - 667,96 ; 82,74 \leq x_4 < 84 \\ -24,64x_4 + 2072,70 ; 84 \leq x_4 < 84,62 \\ 2,36x_4 - 212,82 ; x_4 \geq 84,62 \end{cases}$$

Interpretasi dari model diatas adalah pada saat TPAK Kabupaten/Kota bernilai kisaran 82,74 hingga 84 dan TPAK naik sebesar satu persen dengan asumsi variabel prediktor lain tetap, maka IPM naik sebesar 7,98 persen. Kabupaten Tapanuli Utara adalah wilayah yang tergolong dalam interval kedua pada variabel prediktor TPAK dengan status IPM

tinggi. Semakin tinggi TPAK, berarti semakin besar pula keterlibatan penduduk usia 15 tahun ke atas ke dalam pasar kerja. TPAK Kabupaten Tapanuli Utara tahun 2015 adalah sebesar 83,57 persen dari seluruh penduduk umur 15 tahun keatas. Berdasarkan data yang diperoleh dari BPS Kabupaten Tapanuli Utara, apabila dilihat menurut jenis kelamin TPAK laki-laki hampir sama dibanding TPAK perempuan yaitu masing-masing 86,35 persen dan 80,97 persen. TPAK perempuan di masa depan, diduga mempunyai potensi untuk meningkat di masa yang akan datang. Hal ini diakibatkan oleh pemahaman kaum perempuan tentang persamaan hak dan kewajiban, di samping adanya ruang bagi kaum perempuan tidak hanya terbatas bekerja di rumah tangga saja tetapi juga layak bekerja serta membina karir untuk memberikan kontribusi pendapatan terhadap dirinya dan atau keluarganya. Selanjutnya apabila suatu Kabupaten/Kota dengan TPAK berkisar 84 hingga 84,62 dan naik satu persen dengan asumsi variabel prediktor lain tetap, maka IPM akan turun sebesar 24,64 persen. Tidak terdapat wilayah yang tergolong ke dalam interval ini. Dalam kondisi TPAK suatu Kabupaten/Kota bernilai lebih dari sama dengan 84,62 dan pada kondisi ini TPAK naik sebesar satu persen, maka IPM naik sebesar 2,36 persen. Wilayah yang tergolong dalam interval keempat pada variabel prediktor TPAK antara lain Kabupaten Nias, Kabupaten Dairi, Kabupaten Karo, Kabupaten Humbang Hasundutan, Kabupaten Pakpak Bharat, Kabupaten Samosir dan Kabupaten Nias Barat. Segmen keempat ini didominasi oleh Kabupaten dengan status IPM sedang yang berarti bahwa wilayah tersebut memiliki IPM berkisar antara 60 hingga 70. Kesamaan dari wilayah yang tergolong dalam interval keempat ini adalah TPAK paling banyak menyerap tenaga kerja di sektor pertanian. Hal yang menyebabkan wilayah tersebut berpotensi pada sektor pertanian baik pertanian tanaman pangan, peternakan, perkebunan, kehutanan dan perikanan adalah lahan yang cukup luas dan subur.

5. Hubungan antara TPT terhadap IPM dengan asumsi variabel prediktor lain tetap, maka model yang didapatkan adalah sebagai berikut.

$$y = 0,95x_5 - 65,21(x_5 - 9,37)_+^1 + 179,58(x_5 - 9,82)_+^1 - 113,81(x_5 - 10,04)_+^1$$

$$= \begin{cases} 0,95x_5 ; x_5 < 9,37 \\ -64,25x_5 + 611,07 ; 9,37 \leq x_5 < 9,82 \\ 115,33x_5 - 1152,49 ; 9,82 \leq x_5 < 10,04 \\ 1,51x_5 - 9,76 ; x_5 \geq 10,04 \end{cases}$$

Interpretasi dari model diatas adalah pada kondisi TPT suatu Kabupaten/Kota bernilai kurang dari 9,37 dan TPT naik sebesar satu persen dengan asumsi variabel prediktor lain tetap, maka IPM akan naik sebesar 0,95 persen. Terdapat beberapa wilayah yang tergolong dalam interval ini dapat disajikan pada Tabel 4.13 berikut.

Tabel 4.13 Daftar Kabupaten/Kota dalam Interval Pertama x_5

No	Kabupaten/Kota
1	Kab. Nias
2	Kab. Mandailing Natal
3	Kab. Tapanuli Selatan
4	Kab. Tapanuli Tengah
5	Kab. Tapanuli Utara
6	Kab. Toba Samosir
7	Kab. Asahan
8	Kab. Simalungun
9	Kab. Dairi
10	Kab. Karo
11	Kab. Deli Serdang
12	Kab. Langkat
13	Kab. Nias Selatan
14	Kab. Humbang Hasundutan
15	Kab. Pakpak Bharat
16	Kab. Samosir
17	Kab. Serdang Bedagai
18	Kab. Batu Bara
19	Kab. Padang Lawas Utara
20	Kab. Padang Lawas

Tabel 4.13 Daftar Kabupaten/Kota dalam Interval Pertama x_5 (Lanjutan)

No	Kabupaten/Kota
21	Kab. Labuhanbatu Selatan
22	Kab. Labuanbatu Utara
23	Kab. Nias Utara
24	Kab. Nias Barat
25	Kota Padangsidimpuan

6. Hubungan antara Angka Harapan Hidup (AHH) terhadap IPM di Provinsi Sumatera Utara dengan asumsi variabel prediktor lain tetap, maka model yang diperoleh adalah sebagai berikut.

$$y = 0,38x_6 + 41,67(x_6 - 70,32)_+^1 - 111,85(x_6 - 70,76)_+^1 + 72,81(x_6 - 70,97)_+^1$$

$$= \begin{cases} 0,38x_6 ; x_6 < 70,32 \\ 42,05x_6 - 2930,45 ; 70,32 \leq x_6 < 70,76 \\ -69,7968x_6 + 4984,20 ; 70,76 \leq x_6 < 70,97 \\ 3,02x_6 - 183,68 ; x_6 \geq 70,97 \end{cases}$$

Interpretasi dari model diatas adalah pada saat AHH suatu Kabupaten/Kota bernilai kisaran 70,32 hingga 70,76 dan AHH naik sebesar satu persen maka IPM naik sebesar 42,05 persen. Kabupaten Simalungun dan Kabupaten Karo adalah wilayah yang tergolong dalam interval kedua dan kedua wilayah ini menyandang status IPM tinggi. Selanjutnya suatu Kabupaten/Kota dengan AHH berkisar 70,76 hingga 70,97 dan nilainya naik satu persen dengan asumsi variabel prediktor lain tetap, maka IPM akan turun sebesar 69,79. Tidak terdapat wilayah yang tergolong ke dalam interval ini. Apabila AHH suatu Kabupaten/Kota bernilai lebih dari sama dengan 70,97 dan pada kondisi ini AHH naik sebesar satu persen, maka IPM naik sebesar 3,02 persen. Wilayah yang tergolong dalam interval keempat pada variabel prediktor AHH antara lain Kabupaten Deli Serdang, Kota Pematangsiantar, Kota Medan dan Kota Binjai. Wilayah tersebut merupakan Kabupaten/Kota yang menyandang status IPM tinggi yang berarti bahwa wilayah tersebut memiliki IPM berkisar antara 70 hingga 80.

BAB V KESIMPULAN DAN SARAN

Bab ini memuat kesimpulan dari hasil analisis data dan pembahasan yang telah dilakukan pada bab IV serta saran yang diberikan penulis agar penelitian selanjutnya dapat lebih baik lagi.

5.1 Kesimpulan

Berdasarkan analisis yang dilakukan pada bab sebelumnya, maka dapat disimpulkan adalah sebagai berikut.

1. Terdapat 4 Kabupaten yang terkategori sebagai wilayah dengan status IPM rendah antara lain Kabupaten Nias, Kabupaten Nias Utara, Kabupaten Nias Selatan dan Kabupaten Nias Barat. Terdapat 17 Kabupaten/Kota terkategori sebagai wilayah yang memiliki status IPM sedang, 12 Kabupaten/Kota terkategori sebagai wilayah dengan status IPM tinggi yaitu angka IPM antara 70 hingga 80 dan tidak terdapat Kabupaten/Kota yang terkategori dalam wilayah dengan status IPM sangat tinggi ($IPM \geq 80$).
2. Model regresi nonparametrik Spline terbaik untuk memodelkan Indeks Pembangunan Manusia di Provinsi Sumatera Utara dengan menggunakan kombinasi titik knot 2,1,2,3,3,3. Seluruh variabel yang digunakan dalam penelitian ini berpengaruh signifikan terhadap IPM Sumatera Utara. Berikut ini adalah model akhir yang diperoleh.

$$y = 0,009 + 0,485x_1 - 0,396(x_1 - 88,97)_+^1 - 7,358 \times 10^{-10}(x_1 - 100)_+^1 - 0,279x_2 + 1,215(x_2 - 83,47)_+^1 + 0,342x_3 - 0,129(x_3 - 47,55)_+^1 + 1,787 \times 10^{-10}(x_3 - 87,53)_+^1 - 0,088x_4 + 8,0731(x_4 - 82,74)_+^1 - 32,6271(x_4 - 84)_+^1 + 27,009(x_4 - 84,62)_+^1 + 0,958x_5 - 65,261(x_5 - 9,37)_+^1 + 179,589(x_5 - 9,82)_+^1 - 113,817(x_5 - 10,04)_+^1 + 0,382x_6 + 41,673(x_6 - 70,32)_+^1 - 11,852(x_6 - 70,76)_+^1 + 72,817(x_6 - 70,97)_+^1$$

dengan keterangan

X_1 : Angka Partisipasi Murni SD

X_2 : Angka Partisipasi Murni SMP

X_3 : Angka Partisipasi Murni SMA

X_4 : Tingkat Partisipasi Angkatan Kerja

X_5 : Tingkat Pengangguran Terbuka

X_6 : Angka Harapan Hidup

Koefisien determinasi yang dihasilkan oleh model terbaik adalah sebesar 91,80 persen yang berarti bahwa variabel prediktor mampu menjelaskan variasi variabel respon yaitu IPM Sumatera Utara sebesar 91,80 persen, sedangkan sisanya koefisien determinasi 8,20 persen lainnya dijelaskan oleh variabel prediktor lain yang tidak termasuk ke dalam model.

5.2 Saran

Saran yang dapat direkomendasikan terhadap peneliti selanjutnya dan pembaca adalah sebagai berikut.

1. Penelitian selanjutnya dapat menambahkan variabel-variabel yang diduga memengaruhi IPM Sumatera Utara.
2. Penelitian ini masih terbatas metode yang digunakan hanya menggunakan metode regresi nonparametrik Spline menggunakan tiga knot, untuk selanjutnya diharapkan bisa menggunakan lebih dari tiga knot.
3. Upaya yang perlu dilakukan pemerintah Sumatera Utara yaitu dengan membuat program-program yang berkaitan dengan kesehatan, pendidikan dan mendapatkan standar hidup yang layak. Selain itu, pemerintah sebaiknya juga memerhatikan variabel yang mempunyai pengaruh cukup besar terhadap IPM seperti TPAK, TPT dan AHH, sehingga nantinya dapat meningkatkan nilai IPM Sumatera Utara.

DAFTAR PUSTAKA

- Agresti, A. (1990). *Categorical Data Analysis*. Canada: John Wiley and Son, Inc.
- Ardhanacitri, D. (2013). *Pemodelan dan Pemetaan Pendidikan di Provinsi Jawa Timur Menggunakan Geographically Weighted Regression*. Tugas Akhir, Institut Teknologi Sepuluh Nopember, Jurusan Statistika, Surabaya.
- Badan Perencanaan Pembangunan Kabupaten Malang. (2014). *Indeks Pembangunan Manusia*. Kabupaten Malang: Badan Perencanaan Pembangunan Kabupaten Malang.
- Badan Pusat Statistik. (2016). *Indeks Pembangunan Manusia 2015*. Jakarta: Badan Pusat Statistik.
- Badan Pusat Statistik Provinsi Sumatera Utara. (2015). *Indikator Kesejahteraan Rakyat Provinsi Sumatera Utara*. Sumatera Utara: Badan Pusat Statistik Provinsi Sumatera Utara.
- Badan Pusat Statistik Provinsi Sumatera Utara. (2016). *Sumatera Utara dalam Angka 2016*. Sumatera Utara: Badan Pusat Statistik Provinsi Sumatera Utara.
- Badan Pusat Statistik Provinsi Sumatera Utara. (2016). *Indeks Pembangunan Manusia Metode Baru*. Sumatera Utara: Badan Pusat Statistik Provinsi Sumatera Utara.
- Budiantara, I. N. (2001). Estimasi Parametrik dan Nonparametrik untuk Pendekatan Kurva Regresi. *Pembicara Utama pada Seminar Nasional Statistika V*, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.
- Budiantara, I. N. (2006). Model Spline dengan Knot Optimum. *Jurnal Ilmu Dasar*, FMIPA, Universitas Jember, 7, 77-85.
- Budiantara, I. N. (2007). Model Keluarga Spline Polinomial Truncated dalam Regresi Semiparametrik. *Jurnal Matematika, Ilmu Pengetahuan Alam dan Pengajarannya Universitas Negeri Malang*, 36(1), 1-16.
- Budiantara, I. N. (2009). Spline dalam Regresi Nonparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang. *Pidato Pengukuhan Untuk Jabatan Guru*

- Besar dalam Bidang Matematika Statistika dan Probabilitas, Jurusan Statistika, Fakultas MIPA, Institut Teknologi Sepuluh Nopember.* Surabaya: ITS Press.
- Draper, N. R., dan Smith, H. (1992). *Applied Regression Analysis, Second Edition.* New York: John Willey & Sons, Inc.
- Eubank, R. L. (1988). *Spline Smoothing and Nonparametric Regression.* New York: Marcel Dekker.Inc.
- Gujarati, D. (1992). *Essential of Econometrics.* New York: McGraw-Hill.Inc.
- Gujarati, D. (2004). *Basic Economics fourth edition.* New York: McGraw-Hill.Inc.
- Hardle, W. (1990). *Applied Nonparametric Regression.* New York: Cambridge University Press.
- Irawan, A., dan Suparmoko, M. (2008). *Ekonomika Pembangunan. Edisi Keenam.* Yogyakarta: BPFE.
- Mantra, I. B. (2003). *Demografi Umum Edisi Kedua.* Yogyakarta: Pustaka Pelajar.
- Melliana, A. (2013). *Analisis Statistika Faktor Yang Mempengaruhi Indeks Pembangunan Manusia di Kabupaten/Kota Provinsi Jawa Timur Menggunakan Regresi Panel.* Tugas Akhir, Institut Teknologi Sepuluh Nopember, Jurusan Statistika, Surabaya.
- Retno, A. T. (2014). *Faktor-Faktor yang Mempengaruhi Indeks Pembangunan Manusia di Jawa Timur dengan Pendekatan Regresi Semiparametrik Spline.* Tugas Akhir, Institut Teknologi Sepuluh Nopember, Jurusan Statistika, Surabaya.
- Sukirno, S. (2004). *Makro Ekonomi Teori Pengantar.* Jakarta: PT. Raja Grafindo Perkasa.
- Todaro, M. P. (2000). *Pembangunan Ekonomi di Dunia Ketiga.* Surabaya: Erlangga.
- Wahda, G. (1990). *Spline Models for Observational Data.* Madison: University of Winsconsin.
- Walpole, R. E. (1995). *Pengantar Statistika Edisi ke-3.* Alih Bahasa: Bambang Sumantri. Jakarta: PT. Gramedia Utama.

- Wei, W. W. (2006). *Time Series Analysis: Univariate and Multivariate Methods (2nd ed.)*. Pearson Addison Wesley.
- Yanthi, N. D. (2016). *Pemodelan Faktor-Faktor yang Memengaruhi Indeks Pembangunan Manusia Menggunakan Regresi Nonparametrik Spline di Jawa Tengah*. Tugas Akhir, Institut Teknologi Sepuluh Nopember, Jurusan Statistika, Surabaya.

(Halaman ini sengaja dikosongkan)

LAMPIRAN

Lampiran 1. Data Indeks Pembangunan Manusia Sumatera Utara dan Faktor-Faktor yang Diduga Memengaruhinya.

No	Kabupaten/Kota	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆
1	Kab. Nias	58.85	96.42	73.55	47.55	85.72	0.92	68.97
2	Kab. Mandailing Natal	63.99	99.29	80.50	64.68	73.09	5.78	61.58
3	Kab. Tapanuli Selatan	67.63	99.40	83.89	75.46	76.87	5.91	63.74
4	Kab. Tapanuli Tengah	67.06	98.51	68.82	62.00	75.94	4.98	66.59
5	Kab. Tapanuli Utara	71.32	100.00	80.19	74.59	83.57	2.56	67.55
6	Kab. Toba Samosir	73.4	99.47	69.94	75.29	80.28	3.47	69.14
7	Kab. Labuhan Batu	70.23	97.88	84.20	61.60	59.95	11.39	69.36
8	Kab. Asahan	68.4	99.59	78.73	65.74	61.16	5.82	67.37
9	Kab. Simalungun	71.24	93.28	87.57	69.21	70.23	5.75	70.34
10	Kab. Dairi	69	99.25	91.42	71.67	86.45	1.26	67.78
11	Kab. Karo	72.69	98.47	81.17	69.94	85.25	2.23	70.62
12	Kab. Deli Serdang	72.79	98.18	79.95	69.57	61.23	6.38	71.00
13	Kab. Langkat	68.53	94.22	78.05	52.76	64.51	8.02	67.63
14	Kab. Nias Selatan	58.74	98.85	69.37	52.55	78.88	0.4	67.66
15	Kab. Humbang Hasundutan	66.03	99.32	73.76	72.28	87.1	1.22	68.10
16	Kab. Pakpak Barat	65.53	99.55	74.08	74.05	87.76	2.88	64.85
17	Kab. Samosir	68.43	92.06	80.27	76.60	88.38	1.28	70.26
18	Kab. Serdang Bedagai	68.01	99.11	82.24	57.39	63.05	7.18	67.47
19	Kab. Batu Bara	66.02	98.70	78.43	55.96	57.72	6.32	65.80

Lampiran 1. Data Indeks Pembangunan Manusia Sumatera Utara dan Faktor-Faktor yang Diduga Memengaruhinya (Lanjutan).

No	Kabupaten/Kota	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆
20	Kab. Padang Lawas Utara	67.35	97.46	77.78	68.70	77.6	5.01	66.50
21	Kab. Padang Lawas	65.99	99.29	81.26	58.83	71.87	5.95	66.31
22	Kab. Labuhanbatu Selatan	69.67	99.82	74.28	73.71	67.73	4.15	68.09
23	Kab. Labuanbatu Utara	69.69	98.71	75.37	61.57	66.94	8.75	68.70
24	Kab. Nias Utara	59.88	96.89	78.34	53.93	79.13	4.02	68.59
25	Kab. Nias Barat	58.25	98.89	78.74	68.42	85.3	2.96	67.94
26	Kota Sibolga	71.64	96.08	73.15	75.35	70.04	10.25	67.70
27	Kota Tanjungbalai	66.74	97.31	72.21	59.12	64.83	10.06	61.90
28	Kota Pematangsiantar	76.34	92.33	70.71	75.25	68.63	9.47	72.29
29	Kota Tebing Tinggi	72.81	94.79	71.18	72.60	62.81	10.46	70.14
30	Kota Medan	78.87	88.97	75.96	71.13	60.28	11.00	72.28
31	Kota Binjai	73.81	92.75	68.52	71.14	65.44	10.00	71.59
32	Kota Padangsidimpuan	72.8	96.74	85.79	87.53	68.27	6.96	68.32
33	Kota Gunungsitoli	66.41	96.05	70.45	71.72	66.5	10.00	70.29

Keterangan:

- Y : Indeks Pembangunan Manusia Provinsi Sumatera Utara
- X₁ : Angka Partisipasi Murni SD
- X₂ : Angka Partisipasi Murni SMP
- X₃ : Angka Partisipasi Murni SMA
- X₄ : Tingkat Partisipasi Angkatan Kerja
- X₅ : Tingkat Pengangguran Terbuka
- X₆ : Angka Harapan Hidup

Lampiran 2. Program Pemilihan Titik Knot Optimal dengan Satu Titik Knot Menggunakan *Software R*

```
GCV1=function(para)
{
data=read.table("d://data.txt",header=FALSE)
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 (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot1[j,i]=a[j]
}
}
a1=length(knot1[,1])
knot1=knot1[2:(a1-1),]
aa=rep(1,p)
data1=matrix(ncol=m,nrow=p)
data2=data[,2:q]
a2=nrow(knot1)
GCV=rep(NA,a2)
Rsqr=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 linear 1 knot", "\n")
cat("=====", "\n")
print(knot1)
cat("=====", "\n")
cat("Rsqr dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print(Rsqr)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print(GCV)
s1=min(GCV)
print(max(Rsqr))
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")

```

```
write.csv(GCV,file="d://file//validasi_GCV26.csv")
write.csv(Rsq,file="d://file//validasi_Rsq26.csv")
write.csv(knot1,file="d://file//validasi_knot26.csv")
}
```

Lampiran 3. Program Pemilihan Titik Knot Optimal dengan Dua Titik Knot Menggunakan *Software R*

```
GCV2=function(para)
{
data=read.table("d://data.txt",header=FALSE)
data=as.matrix(data)
p=length(data[,1])
q=length(data[1,])
m=ncol(data)-1-para
dataA=data[, (para+2):q]
F=matrix(0,nrow=p,ncol=p)
diag(F)=1
nk= length(seq(min(data[,para+2]),max(data[,para+2]),length.out=50))
knot=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot[j,i]=a[j]
}
}
a1=nrow(knot)
knot=knot[2:(a1-1),]
a2=nrow(knot)
z=(a2*(a2-1)/2)
knot2=cbind(rep(NA,(z+1)))
for (i in (1:m))
{
knot1=rbind(rep(NA,2))
for ( j in 1:(a2-1))
{
for (k in (j+1):a2)
{
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[, (para+2):q]
data3=data[,2:q]
a3=length(knot2[,1])
GCV=rep(NA,a3)
Rsqr=rep(NA,a3)
for (i in 1:a3)
{
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,data3,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
}
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 linear 2 knot","\n")
cat("=====\n")
print (knot2)
cat("=====\n")
cat("Rsq dengan Spline linear 2 knot","\n")
cat("=====\n")
print (Rsq)
cat("=====\n")
cat("HASIL GCV dengan Spline linear 2 knot","\n")
cat("=====\n")
print (GCV)
s1=min(GCV)
cat("=====\n")
cat("HASIL GCV terkecil dengan Spline linear 2 knot","\n")
cat("=====\n")
cat(" GCV =",s1,"\n")
write.csv(GCV,file="d://file//validasi_GCV26.csv")
write.csv(Rsq,file="d://file//validasi_Rsq26.csv")
write.csv(knot1,file="d://file//validasi_knot26.csv")
}

```

Lampiran 4. Program Pemilihan Titik Knot Optimal dengan Tiga Titik Knot Menggunakan *Software R*

```

GCV3=function(para)
{
data=read.table("d:/data.txt",header=FALSE)
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 (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot[j,i]=a[j]
}
}
knot=knot[2:(nk-1),]
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)
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")
print (knot1)
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.csv(GCV,file="d://file//validasi_GCV36.csv")
write.csv(Rsq,file="d://file//validasi_Rsq36.csv")
write.csv(knot1,file="d://file//validasi_knot36.csv")
}

```

Lampiran 5. Program Pemilihan Titik Knot Optimal dengan Kombinasi Titik Knot Menggunakan *Software R*

```
GCVkom=function(para)
{
data=read.table("d://data.txt",header=FALSE)
data=as.matrix(data)
p1=length(data[,1])
q1=length(data[1,])
v=para+2
F=matrix(0,nrow=p1,ncol=p1)
diag(F)=1
x1=read.table("d://file//x1.txt")
x2=read.table("d:// file //x2.txt")
x3=read.table("e:// file //x3.txt")
x4=read.table("e:// file//x4.txt")
x5=read.table("e:// file //x5.txt")
x6=read.table("e:// file //x6.txt")
n2=nrow(x1)
a=matrix(nrow=6,ncol=3^6)
m=0
for (ii in 1:3)
for (j in 1:3)
for (k in 1:3)
for (l in 1:3)
for (s in 1:3)
for (ee in 1:3)
{
m=m+1
a[,m]=c(ii,j,k,l,s,ee)
}
a=t(a)
GCV=matrix(nrow=nrow(x1),ncol=3^6)
R=matrix(nrow=nrow(x1),ncol=3^6)
for (i in 1:3^6)
{
for (h in 1:nrow(x1))
{
if (a[i,1]==1)
{
gab=as.matrix(x1[,1])
```

```

gen=as.matrix(data[,v])
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,1]==2)
{
gab=as.matrix(x1[,2:3])
gen=as.matrix(cbind(data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1: nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x1[,4:6])
gen=as.matrix(cbind(data[,v],data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,2]==1)
{
gab=as.matrix(x2[,1] )
gen=as.matrix(data[, (v+1)])
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}

```

```

}
else
if (a[i,2]==2)
{
gab=as.matrix(x2[,2:3] )
gen=as.matrix(cbind(data[, (v+1)], data[, (v+1)]))
bb=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x2[,4:6])
gen=as.matrix(cbind(data[, (v+1)], data[, (v+1)], data[, (v+1)]))
bb=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,3]==1)
{
gab=as.matrix(x3[,1] )
gen=as.matrix(data[, (v+2)])
cc=matrix(nrow=nrow(x1)*nrow(data), ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,3]==2)
{
gab=as.matrix(x3[,2:3] )
gen=as.matrix(cbind(data[, (v+2)], data[, (v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data), ncol=2)

```

```

for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x3[,4:6])
gen=as.matrix(cbind(data[, (v+2)], data[, (v+2)], data[, (v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,4]==1)
{
gab=as.matrix(x4[,1] )
gen=as.matrix(data[, (v+3)])
dd=matrix(nrow=nrow(x1)*nrow(data), ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,4]==2)
{
gab=as.matrix(x4[,2:3] )
gen=as.matrix(cbind(data[, (v+3)], data[, (v+3)]))
dd=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
}
else

```

```

{
gab=as.matrix(x4[,4:6])
gen=as.matrix(cbind(data[, (v+3)], data[, (v+3)], data[, (v+3)]))
dd=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,5]==1)
{
gab=as.matrix(x5[,1] )
gen=as.matrix(data[, (v+4)])
ee=matrix(nrow=nrow(x1)*nrow(data), ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,5]==2)
{
gab=as.matrix(x5[,2:3] )
gen=as.matrix(cbind(data[, (v+4)], data[, (v+4)]))
ee=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x5[,4:6])
gen=as.matrix(cbind(data[, (v+4)], data[, (v+4)], data[, (v+4)]))
ee=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{

```

```

if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,6]==1)
{
gab=as.matrix(x6[,1] )
gen=as.matrix(data[, (v+5)])
ff=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ff[w,j]=0 else ff[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,6]==2)
{
gab=as.matrix(x6[,2:3] )
gen=as.matrix(cbind(data[, (v+5)],data[, (v+5)]))
ff=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ff[w,j]=0 else ff[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x6[,4:6])
gen=as.matrix(cbind(data[, (v+5)],data[, (v+5)],data[, (v+5)]))
ff=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ff[w,j]=0 else ff[w,j]=gen[w,j]-gab[h,j]
}
}
ma=as.matrix(cbind(aa,bb,cc,dd,ee,dd,ff))
mx=cbind(rep(1,nrow(data)), data[,2:q1],na.omit(ma))
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:nrow(data))
{
sum=(data[r,1]-yhat[r,])^2
sum1=(yhat[r,]-mean(data[,1]))^2
SSE=SSE+sum
SSR=SSR+sum1
}
Rsq=(SSR/(SSE+SSR))*100
R[i]=Rsq
MSE=SSE/p1
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p1)^2
GCV[h,i]=MSE/A2
}
if (a[i,1]==1) sp=x1[,1] else
if (a[i,1]==2) sp=x1[,2:3] else
sp=x1[,4:6]
if (a[i,2]==1) spl=x2[,1] else
if (a[i,2]==2) spl=x2[,2:3] else
spl=x2[,4:6]
if (a[i,3]==1) splin=x3[,1] else
if (a[i,3]==2) splin=x3[,2:3] else
splin=x3[,4:6]
if (a[i,4]==1) spline=x4[,1] else
if (a[i,4]==2) spline=x4[,2:3] else
spline=x4[,4:6]
if (a[i,5]==1) splines=x5[,1] else
if (a[i,5]==2) splines=x5[,2:3] else
splines=x5[,4:6]
if (a[i,6]==1) spline6=x6[,1] else
if (a[i,6]==2) spline6=x6[,2:3] else
spline6=x6[,4:6]
kkk=cbind(sp,spl,splin,spline,splines,spline6)
cat("=====", "\n")
print(i)
print(kkk)

```

```
print(Rsq)
}
write.csv(GCV,file="d://file//output GCVkom67.csv")
}
```

Lampiran 6. Program Estimasi Parameter dengan Kombinasi Titik Knot 2,1,2,3,3,3

```

uji=function(alpha,para)
{
data=read.table("d://data.txt",header=FALSE)
knot=read.table("d:// knot.txt",header=FALSE)
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+1],data[,m+2],data[,m+2],data[,m+3],d
ata[,m+3],data[,m+3],data[,m+4],data[,m+4],data[,m+4]
,data[,m+5],data[,m+5],data[,m+5])
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.knot[,1:2],data[,3],data.knot[,3],data[,4],data.knot[,4:5],data[,5],da
ta.knot[,6:8],data[,6],data.knot[,9:11],data[,7],data.knot[,12:14])
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)
Rsqr=(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("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.csv(res,file="d://file//output uji residual.csv")
write.csv(pval,file="d://file//output uji pvalue.csv")
write.csv(mx,file="d://file//output uji mx.csv")
write.csv(yhat,file="d://file//output uji yhat.csv")
write.csv(B,file="d://file//output B.csv")
write.csv(thit,file="d://file//output thit.csv")
}

```

Lampiran 7. Program Uji Glejser untuk Kombinasi Titik Knot 2,1,2,3,3,3

```

glejser=function(alpha,para)
{
data=read.table("d://data.txt")
knot=read.table("d://knot.txt",header=FALSE)
res=read.table("d://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+1],data[,m+2],data[,m+2],data[,m+3],data[,m+3],data[,m+3],data[,m+4],data[,m+4],data[,m+4],data[,m+5],data[,m+5],data[,m+5])
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.knot[,1:2],data[,3],data.knot[,3],data[,4],data.knot[,4:5],data[,5],data.knot[,6:8],data[,6],data.knot[,9:11],data[,7],data.knot[,12:14])
mx=as.matrix(mx)
B=(ginv(t(mx)%*%mx))%*%t(mx)%*%res
n1=nrow(B)
yhat=mx%*%B
SSE=sum((res-yhat)^2)
SSR=sum((yhat-rbar)^2)
SST=SSR+SSE
MSE=SSE/(p-n1)

```

```

MSR=SSR/(n1-1)
Rsqr=(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,"")
cat("Error      ",p-n1," ",SSE,"",MSE,"")
cat("Total      ",p-1," ",SST,"")
cat("=====", "\n")
cat("s=",sqrt(MSE),"    Rsqr=",Rsqr,"")
cat("pvalue(F)=",pvalue,"")
}

```


Lampiran 10. *Output* Nilai GCV dengan Tiga Titik Knot

No	Knot						GCV
	x_1	x_2	x_3	x_4	x_5	x_6	
1	89.195	68.987	48.366	58.346	0.624	61.799	18.956
	89.420	69.455	49.182	58.971	0.849	62.017	
	89.645	69.922	49.998	59.597	1.073	62.236	
2	89.195	68.987	48.366	58.346	0.624	61.799	24.129
	89.420	69.455	49.182	58.971	0.849	62.017	
	89.870	70.389	50.814	60.223	1.297	62.454	
3	89.195	68.987	48.366	58.346	0.624	61.799	26.557
	89.420	69.455	49.182	58.971	0.849	62.017	
	90.096	70.857	51.630	60.849	1.521	62.673	
4	89.195	68.987	48.366	58.346	0.624	61.799	26.719
	89.420	69.455	49.182	58.971	0.849	62.017	
	90.321	71.324	52.446	61.474	1.746	62.891	
5	89.195	68.987	48.366	58.346	0.624	61.799	27.945
	89.420	69.455	49.182	58.971	0.849	62.017	
	90.546	71.791	53.261	62.100	1.970	63.110	
6	89.195	68.987	48.366	58.346	0.624	61.799	28.545
	89.420	69.455	49.182	58.971	0.849	62.017	
	90.771	72.259	54.077	62.726	2.194	63.329	
7	89.195	68.987	48.366	58.346	0.624	61.799	30.184
	89.420	69.455	49.182	58.971	0.849	62.017	
	90.996	72.726	54.893	63.351	2.419	63.547	
8	89.195	68.987	48.366	58.346	0.624	61.799	30.398
	89.420	69.455	49.182	58.971	0.849	62.017	
	91.221	73.193	55.709	63.977	2.643	63.766	

Lampiran 11. Output Estimasi Parameter dan Uji Signifikansi Parameter Model

```
=====
Estimasi Parameter
=====
```

```
      [,1]
[1,] 9.916529e-03
[2,] 4.853274e-01
[3,] -3.969462e-01
[4,] -7.358331e-10
[5,] -2.790865e-01
[6,] 1.215505e+00
[7,] 3.422576e-01
[8,] -1.292733e-01
[9,] 1.787792e-10
[10,] -8.858189e-02
[11,] 8.073154e+00
[12,] -3.262717e+01
[13,] 2.700945e+01
[14,] 9.583010e-01
[15,] -6.526123e+01
[16,] 1.795892e+02
[17,] -1.138170e+02
[18,] 3.822380e-01
[19,] 4.167318e+01
[20,] -1.118521e+02
[21,] 7.281793e+01
```

```
-----
Kesimpulan hasil uji serentak
-----
```

Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan

```
-----
Kesimpulan hasil uji individu
-----
```

Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue

0.3018876
 Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.10408
 Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.493332
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.01639768
 Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.1113019
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.01737014
 Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.1701608
 Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.5424423
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.0068419
 Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.4466956
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.04327906
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.007898937
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.004433069
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.03516052
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.02797945
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.04837677
 Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.07457673
 Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.2591181
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.001880631
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.004205021
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.008710195

=====

nilai t hitung

=====

[,1]

[1,] 1.0787923

[2,] 1.7586894

[3,] -0.7065591

[4,] -2.7882496

[5,] -1.7188779

```

[6,] 2.7571658
[7,] 1.4592939
[8,] -0.6269407
[9,] 3.2589535
[10,] -0.7867365
[11,] 2.2591262
[12,] -3.1815159
[13,] 3.4936942
[14,] 2.3737518
[15,] -2.4987507
[16,] 2.1972282
[17,] -1.9526941
[18,] 1.1845721
[19,] 3.9637306
[20,] -3.5223733
[21,] 3.1288654

```

Analysis of Variance

```

=====
Sumber      df    SS      MS      Fhit
Regresi     20  691.0807  34.55403  6.723995
Error       12   61.66697  5.138914
Total       32  752.7476
=====
s= 2.266917 Rsq= 91.80775
pvalue(F)= 0.0007891619

```

Lampiran 12. *Output Uji Glejser*

Kesimpulan hasil uji serentak

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

Analysis of Variance
 =====

Sumber	df	SS	MS	Fhit
Regresi	20	15.74475	0.7872377	0.8192158
Error	12	11.53158	0.960965	
Total	32	27.27633		

=====

s= 0.9802882 Rsq= 57.72313

pvalue(F)= 0.6650481

Lampiran 13. Surat Pernyataan Data Sekunder**SURAT PERNYATAAN**

Saya yang bertanda tangan di bawah ini, mahasiswa Jurusan Statistika FMIPA ITS:

Nama : Nicea Roona Paranoan
NRP : 1313 100 125

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 : Website Badan Pusat Statistik Provinsi Sumatera Utara

Keterangan:

1. Provinsi Sumatera Utara Dalam Angka 2016
2. Profil Kependudukan Provinsi Sumatera Utara Hasil SUPAS 2015

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

Mengetahui
Pembimbing Tugas Akhir

Surabaya, 7 Juli 2017

(Prof. Dr. I Nyoman Budiantara, M.Si)
NIP. 19650603 198903 1 003

(Nicea Roona Paranoan)
NRP 1313 100 125

(Halaman ini sengaja dikosongkan)

BIODATA PENULIS



Nicea Roona Paranoan atau yang akrab disapa Nicea merupakan anak kedua dari pasangan Pasang Paranoan dan Yosпина Kamma dari tiga bersaudara yang lahir di Surabaya pada tanggal 12 Maret 1995. Penulis berdomisili di Surabaya dan telah menempuh pendidikan formal di SDN3 Surodakan (2001-2007), SMPN 1 Trenggalek (2007-2009), SMPN 12 Surabaya (2009-2010), SMAN 10 Surabaya (2010-2013), hingga melanjutkan pendidikan ke jenjang perguruan tinggi di Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember Surabaya dengan jalur Mandiri pada tahun 2013. Semasa kuliah yang ditempuh dalam 4 tahun, penulis aktif di organisasi kemahasiswaan ITS tingkat jurusan yakni Professional Statistics HIMASTA-ITS pada periode 2015-2016 sebagai Manager Operasional. Penulis juga turut berpartisipasi dalam kepanitiaan dilingkup di Jurusan Statistika ITS yaitu panitia Statistics Competition dan Cerdas Bersama Statistika (CERITA).

Segala kritik dan saran serta diskusi lebih lanjut mengenai Tugas Akhir ini dapat dikirimkan melalui surat elektronik (*e-mail*) ke nicea13@mhs.statistika.its.ac.id.

(Halaman ini sengaja dikosongkan)