

**PENERAPAN METODE *MULTIVARIATE ADAPTIVE REGRESSION
SPLINES* (MARS) DALAM PEMODELAN PENDAPATAN DAERAH
DI PROVINSI LAMPUNG**

(SKRIPSI)

Oleh

**KATARINA VANI WULANDARI
NPM. 2217031111**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG**

2026

ABSTRACT

APPLICATION OF THE MULTIVARIATE ADAPTIVE REGRESSION SPLINES (MARS) METHOD IN REGIONAL INCOME MODELING IN LAMPUNG PROVINCE

By

KATARINA VANI WULANDARI

Multivariate Adaptive Regression Spline (MARS) is a development of nonparametric regression used to model the relationship between dependent variables and several independent variables, especially in dealing with nonlinear and high-dimensional data problems. This study applies the MARS method to model regional income of districts/cities in Lampung Province with the aim of obtaining the best model and identifying factors that influence regional income. The data used are secondary data for the period 2015-2024 with predictor variables population density (X_1), Gross Regional Domestic Product/GRDP (X_2), Human Development Index/HDI (X_3) and Poor Population (X_4). Modeling is done by gradually forming basis function, knot point selection, interaction between variables, and model pruning to obtain the optimal model combination. The results of the study show that the best MARS model is able to explain variations in regional income well, where the most influential variable is the poor population with an importance value of 100%, followed by GRDP of 45.4%, HDI of 27.3%, and population density of 22.3%.

Keywords: MARS, Nonparametric, Basis Function.

ABSTRAK

PENERAPAN METODE *MULTIVARIATE ADAPTIVE REGRESSION SPLINES* (MARS) DALAM PEMODELAN PENDAPATAN DAERAH DI PROVINSI LAMPUNG

Oleh

KATARINA VANI WULANDARI

Multivariate Adaptive Regression Spline (MARS) merupakan pengembangan dari regresi nonparametrik yang digunakan untuk memodelkan hubungan antara variabel dependen dengan beberapa variabel independen, khususnya dalam menangani permasalahan data yang bersifat nonlinier dan berdimensi tinggi. Penelitian ini menerapkan metode MARS untuk memodelkan pendapatan daerah kabupaten/kota di Provinsi Lampung dengan tujuan memperoleh model terbaik serta mengidentifikasi faktor-faktor yang mempengaruhi pendapatan daerah. Data yang digunakan merupakan data sekunder periode tahun 2015-2024 dengan variabel prediktor meliputi kepadatan penduduk (X_1), Produk Domestik Regional Bruto/PDRB (X_2), Indeks Pembangunan Manusia/IPM (X_3) dan Penduduk Miskin (X_4). Pemodelan dilakukan melalui pembentukan *basis function* secara bertahap, pemilihan titik knot, interaksi antarvariabel, serta pemangkasan model untuk mendapatkan kombinasi model optimal. Hasil penelitian menunjukkan bahwa model MARS terbaik mampu menjelaskan variasi pendapatan daerah dengan baik, di mana variabel yang paling berpengaruh adalah penduduk miskin dengan nilai kepentingan sebesar 100%, diikuti oleh PDRB sebesar 45,4%, IPM sebesar 27,3%, dan kepadatan penduduk sebesar 22,3%.

Kata-kata kunci: MARS, Nonparametrik, Fungsi Basis.

**PENERAPAN METODE *MULTIVARIATE ADAPTIVE REGRESSION
SPLINES* (MARS) DALAM PEMODELAN PENDAPATAN DAERAH
DI PROVINSI LAMPUNG**

KATARINA VANI WULANDARI

Skripsi

Sebagai Salah Satu Syarat untuk Memperoleh Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika

Fakultas Matematika dan Ilmu Pengetahuan Alam



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG**

2026

Judul Skripsi : **PENERAPAN METODE *MULTIVARIATE ADAPTIVE REGRESSION SPLINES* DALAM PEMODELAN PENDAPATAN DAERAH DI PROVINSI LAMPUNG**

Nama Mahasiswa : **Katarina Vani Wulandari**


Nomor Pokok Mahasiswa : **2217031111**

Program Studi : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**



1. Komisi Pembimbing

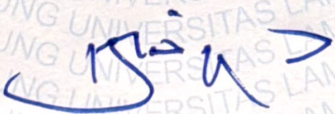

Widiarti, S.Si., M.Si.

NIP. 198005022005012003


Dr. Edwin Russel, S.E., M.Sc.

NIP. 198406192024061001

2. Wakil Dekan Bidang Akademik dan Kerjasama
FMIPA Universitas Lampung


Mulyono, S.Si., M.Si., Ph.D.

NIP. 197406112000031002

MENGESAHKAN

1. Tim Puji

Ketua : Widiarti, S.Si., M.Si.

Sekretaris : Dr. Edwin Russel, S.E., M.Sc.

**Penguji
Bukan Pembimbing : Dr. Bernadhita Herindri
Samodera Utami, S.Si., M.Sc.**

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Heri Satria, S.Si., M.Si.

NIP. 197110012005011002

Tanggal Lulus Ujian Skripsi: 11 Mei 2026

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini:

Nama : **KATARINA VANI WULANDARI**
Nomor Pokok Mahasiswa : **2217031111**
Jurusan : **Matematika**
Judul Skripsi : **PENERAPAN METODE *MULTIVARIATE ADAPTIVE REGRESSION SPLINES* (MARS) DALAM PEMODELAN PENDAPATAN DAERAH DI PROVINSI LAMPUNG**

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri. Apabila dikemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 11 Mei 2026

Penulis.



KATARINA VANI WULANDARI

RIWAYAT HIDUP

Penulis bernama lengkap Katarina Vani Wulandari yang lahir di Way Kanan pada tanggal 28 April 2004. Penulis merupakan anak pertama dari pasangan bapak Petrus Dwi Erwanto dan ibu Lusiana Susanti. Penulis berkebangsaan Indonesia dan beragama Katolik.

Pendidikan yang telah ditempuh penulis dimulai dari Sekolah Dasar (SD) di SDs Bhakti Baradatu dan lulus pada tahun 2016. Selanjutnya penulis melanjutkan pendidikan Sekolah Menengah Pertama (SMP) di SMPs Bhakti Baradatu dan lulus pada tahun 2019. Kemudian melanjutkan pendidikan Sekolah Menengah Atas (SMA) di SMAs Bhakti Baradatu dan lulus pada tahun 2022.

Pada tahun 2022 penulis diterima sebagai mahasiswa di Program Studi S1 Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung melalui jalur SBMPTN. Selama masa perkuliahan, penulis aktif mengikuti kegiatan akademik maupun nonakademik yang mendukung pengembangan kemampuan dan wawasan keilmuan. Pada tahun 2023 penulis juga aktif dalam organisasi, penulis menjadi anggota bidang minat dan bakat Himpunan Mahasiswa Jurusan Matematika (HIMATIKA). Selain itu, penulis sering mengikuti pentas seni maupun perlombaan tari mewakili jurusan maupun fakultas.

KATA INSPIRASI

“Orang lain tidak akan paham *struggle* dan masa sulitnya kita, yang mereka ingin tahu hanya bagian *success stories*. Berjuanglah untuk diri sendiri walaupun tidak ada yang bertepuk tangan, kelak diri kita di masa depan akan sangat bangga dengan apa yang kita perjuangkan hari ini, tetap berjuang ya!”

“Sebab itu janganlah kamu khawatir akan hari esok, karena hari esok mempunyai kesusahannya sendiri. Kesusahan sehari cukuplah untuk sehari.”

(Matius 6:34)

“Setiap tujuan pasti ada prosesnya, setiap prosesnya pasti ada badainya, bertahanlah Tuhan pasti menolong.”

(Filipi 4:6)

“Kamu tidak akan gagal, kamu akan berhasil sesuai dengan rencana Tuhan.”

(Yesaya 55:8–9)

“Pada akhirnya, ini semua hanyalah permulaan.”

(Nadin Amizah)

“Tugasku hanya bertahan, terus jalan dan mengalirlah seperti air. Dari lahir sampai ku jadi debu di akhir.”

(Donne Maulana)

PERSEMBAHAN

Skripsi ini dipersembahkan sebagai ungkapan syukur kepada Tuhan Yang Maha Esa atas limpahan rahmat dan pertolongan-Nya kepada penulis sehingga skripsi ini dapat terselesaikan dengan baik. Dengan penuh rasa syukur karya sederhana ini penulis persembahkan untuk:

Kedua Orang Tuaku

Teruntuk cinta pertama dan panutan penulis, Bapak Petrus Dwi Erwanto. Beliau memang tidak sempat merasakan pendidikan sampai bangku perkuliahan, namun beliau mampu mendidik penulis, memotivasi, memberikan arahan, usaha dan dukungan hingga penulis mampu menyelesaikan studinya sampai sarjana. Teruntuk Mamak tercinta, sosok yang luar biasa yang selalu hadir dalam doa, semangat, dan pengorbanan tanpa henti dan tanpa mengeluh. Terima kasih atas segala cinta yang takkan tergantikan, atas air mata dan doa yang selalu menyertai setiap langkah penulis hingga sampai di titik ini. Semoga Tuhan senantiasa memberikan kesehatan, umur yang panjang, dan kebahagiaan kepada Bapak dan Mamak.

Dosen Pembimbing

Terima kasih kepada dosen pembimbing dan pembahas yang telah membantu memberikan motivasi, kritik, saran, serta ilmu yang sangat bermanfaat.

Almamater Tercinta Universitas Lampung

SANWACANA

Puji dan syukur penulis panjatkan kehadirat Tuhan Yang Maha Esa, atas segala berkat dan penyertaannya-Nya sehingga penulis dapat menyelesaikan Skripsi dengan judul “Penerapan Metode *Multivariate Adaptive Regression Splines* (MARS) Dalam Pemodelan Pendapatan Daerah Di Provinsi Lampung”.

Dalam menyusun skripsi ini, penulis banyak mendapatkan dukungan, bimbingan dan bantuan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada:

1. Ibu Widiarti, S.Si., M.Si., selaku dosen pembimbing satu yang senantiasa memberikan bimbingan, arahan, bantuan, dan saran kepada penulis dalam menyusun skripsi ini.
2. Bapak Dr. Edwin Russel, S.E., M.Sc., selaku pembimbing dua yang telah memberikan bimbingan, pengarahan dan saran yang bermanfaat kepada penulis dalam menyusun skripsi ini.
3. Dr. Bernadhita Herindri Samodera Utami, S.Si., M.Sc., selaku dosen pembahas yang telah memberikan masukan, kritik dan saran kepada penulis dalam menyusun skripsi ini.
4. Ibu Siti Laelatul Chasanah, S.Pd, M.Si., selaku dosen Pembimbing Akademik yang telah membimbing penulis selama masa perkuliahan.
5. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika.
6. Bapak Dr. Eng. Heri Satria, S.Si., M.Si., selaku Dekan Fakultas Matematika.
7. Seluruh dosen, staff dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

8. Kedua orang tua penulis tercinta, Bapak Petrus Dwi Erwanto dan Ibu Lusia Susanti serta adik penulis sayangi Fransiskus Alviano Saputra yang selalu memberikan doa, nasihat, dukungan, memotivasi dan semangat yang menjadi kekuatan terbesar dalam setiap langkah penulis.
9. Seluruh keluarga besar penulis serta keluarga besar Nicolas Augusto Rumahorbo yang senantiasa memberikan semangat, dukungan, motivasi dan arahan kepada penulis selama perkuliahan dan menyusun skripsi ini.
10. Teman-teman penulis Aini, Anggun, Aurel, Fika, Farhan, Kirei dan Leony atas kebersamaan selama masa perkuliahan, serta selalu memberikan semangat, masukan dan hiburan selama penulis menyelesaikan skripsi ini.
11. Seluruh pihak terkait yang tidak dapat disebutkan satu persatu yang telah membantu penulis dalam menyelesaikan skripsi ini.

Penulis menyadari skripsi ini masih jauh dari kata sempurna dan masih terdapat banyak kekurangan baik dalam penyajian maupun teknik penulisan. Oleh sebab itu, saran dan kritikan yang membangun senantiasa penulis harapkan demi menyempurnakan skripsi ini.

Bandar Lampung, 11 Mei 2026

Katarina Vani Wulandari

DAFTAR ISI

DAFTAR ISI	xiii
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xiv
I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Tujuan Penelitian	3
1.3 Manfaat Penelitian	3
II TINJAUAN PUSTAKA	2
2.1 Statistika Deskriptif	2
2.2 Regresi Nonparametrik	2
2.2.1 Uji Linearitas Ramsey RESET	4
2.3 <i>Multivariate Adaptive Regression Splines</i> (MARS)	5
2.3.1 Standardisasi Data Menggunakan Z-Score	8
2.3.2 Pembentukan Model MARS	9
2.3.3 Pendugaan Parameter Model MARS	11
2.4 Pemilihan Model MARS Terbaik	13
2.5 Pendapatan Daerah	14
2.6 Kepadatan Penduduk	14
2.7 Produk Domestik Regional Bruto (PDRB) Per Kapita	15
2.8 Indeks Pembangunan Manusia (IPM)	15
2.9 Penduduk Miskin	16
III METODE PENELITIAN	17
3.1 Waktu dan Tempat Penelitian	17
3.2 Data Penelitian	17
3.3 Metode Penelitian	17
IV HASIL DAN PEMBAHASAN	20
4.1 Eksplorasi Data Pendapatan Daerah Provinsi Lampung	20
4.2 Analisis Pola Hubungan Data	28
4.3 <i>Pemodelan Multivariate Adaptive Regression Splines</i> (MARS)	30

4.4	Pemilihan Model <i>Multivariate Adaptive Regression Splines</i> (MARS) Terbaik	32
4.5	Evaluasi Kinerja dan Tingkat Kepentingan Variabel Model MARS Terbaik	35
V	KESIMPULAN	39
5.1	Kesimpulan	39
	DAFTAR PUSTAKA	40

DAFTAR TABEL

4.1	Statistika Deskriptif Data Penelitian	21
4.2	Hasil Uji Linearitas	30
4.3	Hasil Kombinasi Model MARS	31
4.4	Model MARS dengan Nilai GCV Terkecil	32
4.5	Tingkat Pengaruh Variabel Prediktor	38

DAFTAR GAMBAR

2.1	Ilustrasi Bentuk Minimum Observasi (MO)	11
3.1	Diagram Alir Penelitian	19
4.1	Diagram Garis Rata-rata Pendapatan Daerah di Provinsi Lampung Tahun 2015–2024	23
4.2	Diagram Garis Rata-rata Kepadatan Penduduk di Provinsi Lampung Tahun 2015–2024	24
4.3	Diagram Garis Rata-rata PDRB Per Kapita di Provinsi Lampung Tahun 2015–2024	25
4.4	Diagram Garis Rata-rata IPM di Provinsi Lampung Tahun 2015–2024	26
4.5	Diagram Garis Rata-rata Penduduk Miskin di Provinsi Lampung Tahun 2015–2024	27
4.6	Plot Hubungan Variabel Respon dan Variabel Prediktor	28
4.7	Plot Aktual Model MARS Terbaik	36
4.8	Plot Residual Model MARS Terbaik	37

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Analisis regresi pertama kali dikembangkan oleh Sir Francis Galton di akhir abad ke-19. Model ini digunakan untuk mengetahui pengaruh variabel independen terhadap variabel dependen (Gujarati dkk., 2009) . Salah satu pendekatan yang umum digunakan adalah regresi parametrik, namun metode ini memiliki keterbatasan karena harus memenuhi asumsi-asumsi ketat sehingga kurang mampu menangani data yang berdimensi tinggi. Pada kenyataannya, data penelitian sering kali tidak memenuhi asumsi-asumsi tersebut sehingga hasil estimasi yang diperoleh menjadi kurang optimal. Sebagai alternatif, digunakan regresi nonparametrik yang lebih fleksibel karena tidak memerlukan bentuk fungsi tertentu dan mampu menyesuaikan diri terhadap pola data yang kompleks. Salah satu metode nonparametrik yang digunakan adalah *Multivariate Adaptive Regression Splines* (MARS) sebagai penyempurnaan pendekatan sebelumnya.

Metode MARS merupakan pengembangan dari regresi nonparametrik yang menggabungkan konsep spline dengan *Recursive Partitioning Regression* (RPR) (Friedman, 1991). Metode ini dirancang untuk menangani data dengan dimensi tinggi, khususnya ketika jumlah variabel independen berada pada kisaran $3 \leq n \leq 20$. Keunggulan MARS adalah tidak memerlukan pemenuhan asumsi-asumsi statistik tertentu seperti normalitas, homoskedastisitas, maupun multikolinearitas, sehingga lebih adaptif digunakan pada berbagai jenis data yang kompleks. MARS bekerja dengan mengombinasikan *basis function* (BF) terbaik untuk membentuk model yang fleksibel, akurat, dan efisien dalam menjelaskan hubungan antara variabel independen dan variabel dependen.

Beberapa penelitian menunjukkan bahwa metode MARS efektif digunakan dalam berbagai bidang. Pratiwi (2020) membuktikan bahwa MARS mampu mengidentifikasi faktor-faktor yang memengaruhi tingkat kelahiran di Kota Kendari. Mattalunru dkk. (2022) menunjukkan bahwa MARS dapat menentukan variabel paling berpengaruh terhadap curah hujan di Kota Makassar. Kartini dkk. (2022) juga menemukan bahwa MARS efektif dalam memodelkan faktor risiko kejadian stunting balita di Kabupaten Bojonegoro. Selain itu, Panggabean dkk. (2023) mengungkapkan bahwa sektor pertanian, kehutanan, perikanan, konstruksi, dan industri pengolahan berpengaruh signifikan terhadap PDRB Sumatra Utara berdasarkan model MARS. Zakiyah dkk. (2024) menunjukkan bahwa IHSG dipengaruhi oleh inflasi, nilai tukar, dan jumlah uang beredar sebagai faktor dominan. Berdasarkan berbagai penelitian tersebut, terlihat bahwa metode MARS mampu digunakan secara efektif pada berbagai bidang, baik ekonomi, sosial, maupun lingkungan, karena kemampuannya dalam memodelkan hubungan yang kompleks dan nonlinier antarvariabel.

Pendapatan daerah dalam beberapa tahun terakhir mengalami perubahan yang tidak selalu stabil, sehingga mengindikasikan adanya pengaruh dari berbagai faktor lain. Pemilihan metode MARS didasarkan pada pertimbangan bahwa hubungan antara pendapatan daerah dengan variabel kepadatan penduduk, Produk Domestik Regional Bruto (PDRB), Indeks Pembangunan Manusia (IPM), dan penduduk miskin cenderung bersifat tidak linier. Hal ini berarti kenaikan salah satu variabel independen tidak selalu diikuti oleh kenaikan pendapatan daerah. Variabel-variabel tersebut dipilih karena merepresentasikan faktor demografis, ekonomi, dan sosial yang memengaruhi pendapatan daerah secara tidak langsung di luar komponen utama pembentuk pendapatan daerah, seperti Pendapatan Asli Daerah (PAD), Transfer ke Daerah dan Dana Desa (TKDD), serta lain-lain pendapatan daerah yang sah, yang secara struktural memiliki hubungan linier. Oleh karena itu, metode MARS diharapkan mampu menangkap pola hubungan yang lebih kompleks serta mengidentifikasi faktor yang paling berpengaruh secara adaptif dan fleksibel.

1.2 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut.

1. Membangun model pendapatan daerah di Provinsi Lampung dan memperoleh model terbaik menggunakan metode *Multivariate Adaptive Regression Splines* (MARS).
2. Mengetahui variabel yang paling berpengaruh terhadap pendapatan daerah di Provinsi Lampung.

1.3 Manfaat Penelitian

Manfaat dari penelitian ini adalah untuk memberikan wawasan terkait analisis menggunakan metode *Multivariate Adaptive Regression Splines* (MARS) serta mengetahui tentang faktor-faktor apa saja yang mempengaruhi pendapatan daerah di Provinsi Lampung. Selain itu, penelitian ini dapat menjadi bahan bacaan dan referensi bagi peneliti selanjutnya.

BAB II

TINJAUAN PUSTAKA

2.1 Statistika Deskriptif

Menurut Walpole (1995), statistika deskriptif adalah metode yang berkaitan dengan pengumpulan dan penyajian suatu data sehingga memberikan informasi yang berguna. Statistik deskriptif berfungsi untuk mendeskripsikan atau memberi gambaran terhadap objek yang diteliti melalui data sampel atau populasi (Sugiyono, 2007). Data yang disajikan dalam statistika deskriptif biasanya dalam bentuk ukuran pemusatan data (Kuswanto, 2012).

Statistika deskriptif pada dasarnya bertujuan untuk memberikan gambaran yang jelas mengenai suatu data sehingga informasi yang terkandung di dalamnya lebih mudah dipahami. Menurut Hasan (2001), statistika deskriptif juga bertujuan untuk menyajikan karakteristik umum dari suatu data melalui berbagai jenis seperti tabel, grafik, ukuran pemusatan (mean, median, modus, nilai maksimum, dan minimum), serta ukuran penyebaran data seperti varians, standar deviasi, dan rentang (range).

2.2 Regresi Nonparametrik

Menurut Gujarati dkk. (2009), analisis regresi adalah suatu metode statistika yang umum digunakan untuk meneliti hubungan antara peubah bebas (independen) dengan peubah tak bebas (dependen). Menurut Budiantara (2005), ditinjau dari pola hubungannya regresi terbagi menjadi regresi parametrik, regresi semiparametrik dan nonparametrik.

Regresi nonparametrik mulai dikenal sekitar abad ke-19, tepatnya pada tahun 1857 (Hardle, 1992). Regresi nonparametrik merupakan salah satu pendekatan yang digunakan untuk mengetahui pola hubungan antara variabel independen dan variabel dependen yang tidak diketahui pola datanya. Karena bentuk pola data regresi tidak diketahui, maka regresi nonparametrik dapat membentuk fungsi apapun dalam menduga bentuk kurva regresi baik linear atau nonlinear.(Eubank, 1999). Berikut model regresi nonparametrik secara umum:

$$y_i = f(x_i) + \varepsilon_i, \quad \text{untuk } i = 1, 2, \dots, n \quad (2.1)$$

dengan:

- y_i = variabel respon pada pengamatan ke-i
- $f(x_i)$ = fungsi regresi nonparametrik
- x_i = variabel prediktor pada pengamatan ke-i
- ε_i = nilai *error*/galat
- n = jumlah pengamatan

Terdapat beberapa metode regresi nonparametrik yang sering digunakan dalam analisis data, antara lain regresi kernel, regresi spline, *local polynomial regression* (LOESS), *K-Nearest Neighbor* (KNN) dan lain sebagainya. Metode-metode tersebut digunakan untuk memodelkan hubungan nonlinier tanpa harus menentukan bentuk fungsi tertentu. Sebagai contoh, regresi spline dapat dinyatakan dalam bentuk:

$$f(x) = \beta_0 + \beta_1 x + \sum_{k=1}^K \beta_{k+1} (x - t_k)_+ + \varepsilon_i \quad (2.2)$$

dengan:

- $f(x)$ = fungsi respon (variabel dependen)
- β_0 = konstanta (intersep)
- β_1, β_{k+1} = parameter/koeffisien model
- x = variabel independen
- t_k = titik knot ke- k
- $(x - t_k)_+$ = fungsi *hinge* (*truncated spline*)
- K = jumlah knot
- ε_i = nilai *error* (galat)

2.2.1 Uji Linearitas Ramsey RESET

Uji linearitas merupakan pengujian yang dilakukan untuk mengetahui apakah hubungan antara variabel respon dan variabel prediktor bersifat linear atau nonlinear. Apabila hubungan yang terbentuk tidak linear, maka pendekatan regresi parametrik menjadi kurang sesuai. Salah satu metode yang dapat digunakan untuk mendeteksi ketidaklinieran model adalah uji Ramsey RESET (*Regression Specification Error Test*). Hasil pengujian ini menjadi dasar dalam pemilihan metode regresi nonparametrik, seperti *Multivariate Adaptive Regression Splines* (MARS), yang mampu memodelkan hubungan non-linear secara fleksibel. Hipotesis yang digunakan dalam uji Ramsey RESET adalah sebagai berikut (Ramsey, 1969):

H_0 : hubungan antara variabel respon dan prediktor bersifat linear

H_1 : hubungan antara variabel respon dan prediktor bersifat nonlinear

Statistik uji Ramsey RESET menggunakan statistik F , yaitu:

$$F = \frac{(RSS_R - RSS_{UR})/q}{RSS_{UR}/(n - k)} \quad (2.3)$$

dengan:

RSS_R = *Residual Sum of Squares* model terbatas,

RSS_{UR} = *Residual Sum of Squares* model tidak terbatas,

q = jumlah parameter tambahan,

n = jumlah pengamatan,

k = jumlah parameter pada model.

Kriteria pengujian yang digunakan adalah:

- Jika nilai $p\text{-value} \leq \alpha$, maka H_0 ditolak sehingga terdapat indikasi hubungan nonlinear.
- Jika nilai $p\text{-value} > \alpha$, maka H_0 tidak ditolak sehingga hubungan antar variabel dianggap linear.

Uji Ramsey RESET dapat digunakan sebagai dasar dalam menentukan apakah model regresi linear sudah memadai atau diperlukan pendekatan pemodelan nonlinear, seperti metode *Multivariate Adaptive Regression Splines* (MARS).

2.3 Multivariate Adaptive Regression Splines (MARS)

Multivariate Adaptive Regression Splines (MARS) merupakan suatu metode regresi nonparametrik yang diperkenalkan pertama kali oleh J.H. Friedman pada tahun 1990. metode ini merupakan pengembangan dari metode spline dan *Recursive Partitioning Regression* (RPR) yang memiliki kendala pada model yang tidak kontinu pada knots (titik potong). Proses pembentukan model MARS melalui dua tahap, yaitu algoritma *forward stepwise* dan *backward pruning*. Tahap *forward stepwise* merupakan proses pembentukan model dengan menambahkan *basis function* secara bertahap dimulai dari model paling sederhana. Pada setiap penambahan BF, parameter model diestimasi menggunakan metode *Ordinary Least Squares* (OLS) untuk memperoleh koefisien yang meminimalkan *error*.

Selanjutnya dihitung nilai *Sum of Squared Error* (SSE) untuk mengevaluasi kinerja model dan BF yang dipilih adalah BF yang mampu memberikan penurunan SSE terbesar. Proses ini dilakukan berulang dengan menambahkan BF satu per satu hingga model menjadi semakin kompleks. Sedangkan *backward pruning* merupakan proses penyederhanaan model dengan menghapus *basis function* yang tidak signifikan secara bertahap. Pemilihan model terbaik dilakukan berdasarkan nilai *Generalized Cross Validation* (GCV) minimum, sehingga struktur knot yang dipilih menghasilkan model dengan kinerja optimal.

Model MARS berguna untuk mengatasi permasalahan data berdimensi tinggi dan menghasilkan prediksi variabel dependen yang akurat serta menghasilkan model kontinu dalam knot berdasarkan nilai GCV terkecil. Data berdimensi tinggi adalah data dengan jumlah variabel prediktor sebanyak $3 \leq r \leq 20$ dan sampel data yang berukuran $50 \leq n \leq 1000$, dengan r adalah banyak variabel prediktor dan n adalah ukuran sampel. Persamaan model MARS dituliskan sebagai berikut (Friedman, 1991):

$$f(x) = a_0 + \sum_{m=1}^M a_m \prod_{k=1}^{K_m} [s_{km} (x_{v(k,m)} - t_{km})]_+ \quad (2.4)$$

dengan:

a_0	= intersep
a_m	= koefisien <i>basis function</i> ke- m , $m = 1, 2, \dots, M$
M	= banyaknya <i>basis function</i>
K_m	= tingkat interaksi dalam satu <i>basis function</i>
s_{km}	= arah <i>basis function</i> terhadap titik knot (nilainya ± 1)
$x_{v(k,m)}$	= variabel prediktor
t_{km}	= nilai knot dari variabel prediktor

Parameter dalam model MARS diduga menggunakan metode *Ordinary Least Squares* (OLS). Arah *basis function* dengan nilai $+1$ menunjukkan bahwa fungsi aktif setelah melewati titik knot, yang dinyatakan dengan bentuk $\max(0, x - t)$, sedangkan -1 menunjukkan bahwa fungsi aktif sebelum titik knot, yang dinyatakan dalam bentuk $\max(0, t - x)$. Bentuk $[z]_+$ merupakan fungsi *hinge* yang didefinisikan sebagai $[z]_+$ yaitu bernilai z jika $z > 0$ dan bernilai 0 jika $z \leq 0$.

Fungsi regresi nonparametrik model MARS dapat dinyatakan dalam persamaan 2.5:

$$\begin{aligned} y_i &= f(x) + \varepsilon_i \\ &= a_0 + \sum_{m=1}^M a_m \prod_{k=1}^{K_m} [s_{km} (x_{v(k,m)} - t_{km})]_+ + \varepsilon_i \end{aligned} \quad (2.5)$$

Dengan fungsi knot linear sebagai berikut:

$$(x_{v(k,m)} - t_{km})_+ = \begin{cases} x_{v(k,m)} - t_{km}, & \text{jika } x_{v(k,m)} - t_{km} > 0 \\ 0, & \text{jika } x_{v(k,m)} - t_{km} \leq 0 \end{cases}$$

Persamaan model MARS dapat dituliskan dalam bentuk notasi matriks seperti persamaan 2.6:

$$\mathbf{Y} = \mathbf{B}\mathbf{a} + \boldsymbol{\varepsilon} \quad (2.6)$$

dengan:

$$\begin{aligned} \mathbf{Y} &= (y_1, y_2, \dots, y_n)' \\ \mathbf{a} &= (a_1, a_2, \dots, a_M)' \\ \boldsymbol{\varepsilon} &= (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)' \end{aligned}$$

Matriks *basis function* \mathbf{B} didefinisikan sebagai:

$$\mathbf{B} = \begin{pmatrix} 1 \prod_{k=1}^{K_1} [s_{1m}(x_{v_1(1,m)} - t_{1m})] & \cdots & \prod_{k=1}^{K_M} [s_{Mm}(x_{v_1(M,m)} - t_{Mm})] \\ 1 \prod_{k=1}^{K_1} [s_{1m}(x_{v_2(1,m)} - t_{1m})] & \cdots & \prod_{k=1}^{K_M} [s_{Mm}(x_{v_2(M,m)} - t_{Mm})] \\ \vdots & \vdots & \ddots & \vdots \\ 1 \prod_{k=1}^{K_1} [s_{1m}(x_{v_n(1,m)} - t_{1m})] & \cdots & \prod_{k=1}^{K_M} [s_{Mm}(x_{v_n(M,m)} - t_{Mm})] \end{pmatrix}$$

dengan:

- \mathbf{Y} : vektor variabel respon berukuran $(n \times 1)$,
- \mathbf{B} : matriks *basis function* berordo $(n \times (M + 1))$,
- \mathbf{a} : vektor koefisien regresi berukuran $((M + 1) \times 1)$,
- ϵ : vektor galat berukuran $(n \times 1)$.

Persamaan (2.4) dapat dituliskan dalam bentuk umum MARS berdasarkan interaksi antarvariabel prediktor sebagai berikut:

$$\begin{aligned} f(x_i) = & a_0 + \sum_{m=1}^M a_m [s_{1m}(x_{v_i(1,m)} - t_{1m})]_+ \\ & + \sum_{m=1}^M a_m [s_{1m}(x_{v_i(1,m)} - t_{1m})]_+ [s_{2m}(x_{v_i(2,m)} - t_{2m})]_+ \\ & + \sum_{m=1}^M a_m [s_{1m}(x_{v_i(1,m)} - t_{1m})]_+ [s_{2m}(x_{v_i(2,m)} - t_{2m})]_+ [s_{3m}(x_{v_i(3,m)} - t_{3m})]_+ \\ & + \cdots \end{aligned} \tag{2.7}$$

Persamaan (2.7) dapat diringkas menjadi:

$$f(x) = a_0 + f_i(x_i) + f_{ij}(x_i, x_j) + f_{ijk}(x_i, x_j, x_k) + \cdots \tag{2.8}$$

Pada dasarnya, *basis function* menunjukkan adanya hubungan antara variabel independen dengan variabel dependen. Persamaan (2.7) menunjukkan bahwa penjumlahan pertama meliputi semua *basis function* untuk satu variabel tanpa interaksi, penjumlahan kedua meliputi semua *basis function* untuk interaksi antara dua variabel, penjumlahan ketiga meliputi semua *basis function* untuk interaksi antara tiga variabel, dan seterusnya. Sehingga dapat diketahui bahwa *basis function* merupakan sekumpulan fungsi yang digunakan untuk merepresentasikan informasi yang terdiri atas satu atau lebih variabel termasuk interaksi antar variabel. Bentuk umum model MARS pada persamaan (2.8) dapat dituliskan dalam bentuk linear terhadap *basis function* (Friedman, 1991)

$$f(x) = a_0 + a_1BF_1 + a_2BF_2 + \dots + a_mBF_m \quad (2.9)$$

dengan:

- a_0 : intersep
- a_m : koefisien *basis function* ke- m ; $m = 1, 2, \dots, M$
- M : banyaknya *basis function*
- BF_m : *basis function* ke- m

2.3.1 Standardisasi Data Menggunakan Z-Score

Sebelum dilakukan pembentukan model *Multivariate Adaptive Regression Splines* (MARS), data variabel prediktor terlebih dahulu dilakukan standardisasi menggunakan metode *Z-Score*. Standardisasi dilakukan untuk menyamakan skala antar variabel sehingga variabel dengan rentang nilai yang besar tidak mendominasi proses pemodelan. Metode *Z-Score* mentransformasikan data ke dalam bentuk yang memiliki rata-rata sama dengan nol dan simpangan baku sama dengan satu. Bentuk umum standardisasi *Z-Score* diberikan sebagai berikut (Walpole dkk., 2012):

$$Z_i = \frac{X_i - \bar{X}}{S} \quad (2.10)$$

dengan:

Z_i = nilai hasil standardisasi,

X_i = nilai data ke- i ,

\bar{X} = rata-rata data,

S = simpangan baku data.

Melalui proses standardisasi, setiap variabel prediktor memiliki skala yang seragam sehingga proses pembentukan model MARS dapat dilakukan dengan lebih optimal.

2.3.2 Pembentukan Model MARS

MARS membangun model melalui dua tahap, yaitu *forward selection* dan *backward pruning*. Tahap *forward selection* dilakukan dengan menambahkan *basis function* secara bertahap untuk membentuk model yang kompleks. Sebaliknya tahap *backward pruning* bertujuan menyederhanakan model dengan menghilangkan *basis function* yang dianggap yang kurang relevan sehingga diperoleh model yang lebih tepat dan efisien (Friedman, 1991).

Menurut Nash & Bradford *et al.*, (2001), hal-hal yang harus diperhatikan dalam membuat model MARS adalah sebagai berikut:

1. Knot

Knot merupakan titik pada variabel prediktor di mana kemiringan (*slope*) pada garis regresi mengalami perubahan. Pada setiap titik knot diharapkan fungsi yang terbentuk tetap bersifat kontinu, sehingga tidak terjadi diskontinuitas antara satu bagian fungsi dengan bagian lainnya.

2. Basis Function (BF)

Basis function atau fungsi basis merupakan komponen utama dalam pembentukan model MARS yang merepresentasikan hubungan antara variabel prediktor dan variabel respon pada interval tertentu. Jumlah maksimum fungsi basis yang digunakan umumnya dibatasi sebesar 2-4 kali jumlah variabel prediktor untuk menjaga fleksibilitas model dan menghindari *overfitting* (Friedman, 1991). *Basis function* berbentuk *piecewise* linear dan bersifat

kontinu pada titik knot sehingga mampu menangkap perubahan pola hubungan data secara adaptif. Berikut merupakan bentuk dari BF:

$$BF(x, t) = \max(0, x - t) = \begin{cases} x - t, & \text{jika } x > t, \\ 0, & \text{jika } x \leq t \end{cases}$$

atau

$$BF(x, t) = \max(0, t - x) = \begin{cases} t - x, & \text{jika } x < t, \\ 0, & \text{jika } x \geq t \end{cases}$$

dengan x merupakan variabel prediktor dan t merupakan titik knot. Fungsi \max (maksimum) digunakan untuk mengambil nilai terbesar antara dua nilai, sehingga *basis function* hanya akan bernilai positif ketika kondisi tertentu terpenuhi, dan bernilai nol jika kondisi tersebut tidak terpenuhi.

3. Maksimum Interaksi (MI)

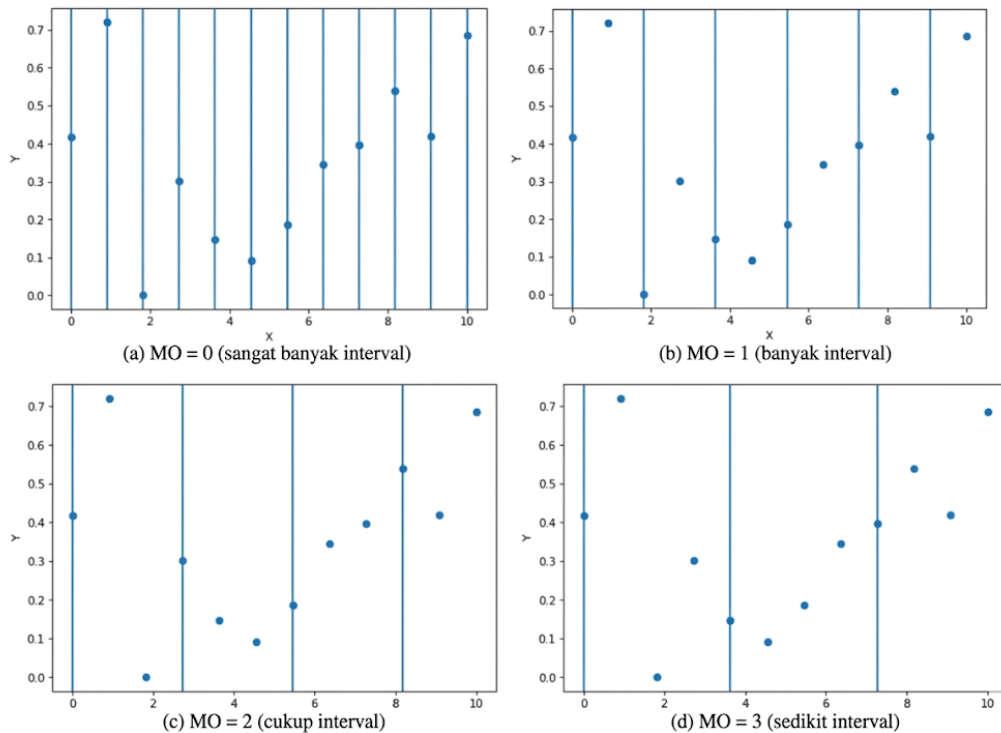
Maksimum interaksi adalah hasil perkalian silang antara dua atau lebih variabel prediktor yang saling berkorelasi. Dalam MARS jumlah Maksimum Interaksi (MI) yang disarankan oleh (Friedman, 1991) adalah 1, 2 atau 3. Jika $MI > 3$ maka akan diperoleh model yang sangat kompleks dan model akan semakin sulit diinterpretasi. Berikut merupakan contoh MI:

- Jika $MI = 1$, model hanya memuat satu interaksi variabel prediktor yaitu dengan bentuk $(x_1 - t_1)_+$.
- Jika $MI = 2$, model memuat dua interaksi variabel prediktor yaitu dengan bentuk $(x_1 - t_1)_+ \times (x_2 - t_2)_+$.
- Jika $MI = 3$, model memuat tiga interaksi variabel prediktor yaitu dengan bentuk $(x_1 - t_1)_+ \times (x_2 - t_2)_+ \times (x_3 - t_3)_+$.

4. Minimum observasi (MO)

Minimum observasi adalah jumlah minimum data yang diizinkan berada pada setiap region atau interval yang terbentuk di antara titik knot. Nilai MO yang disarankan oleh (Friedman, 1991) antara knot adalah 0, 1, 2 dan 3 observasi. MO membantu menjaga model tetap stabil. MO yang terlalu kecil dapat menyebabkan model sangat sensitif terhadap perubahan kecil pada data, sedangkan MO yang terlalu besar dapat menyebabkan hilangnya

pola lokal yang penting. Dengan demikian, pemilihan MO yang sesuai akan memberikan keseimbangan antara fleksibilitas model dan stabilitas estimasi. berikut ilustrasi MO:



Gambar 2.1 Ilustrasi Bentuk Minimum Observasi (MO)

2.3.3 Pendugaan Parameter Model MARS

Pendugaan parameter pada metode *Multivariate Adaptive Regression Splines* (MARS) untuk variabel bersifat kontinu menggunakan metode kuadrat terkecil atau *Ordinary Least Square* (OLS). Metode ini meminimalkan jumlah kuadrat residual, sehingga parameter yang diestimasi memberikan kecocokan model terbaik terhadap data melalui kombinasi *basis function* yang terbentuk dalam MARS (Friedman, 1991). Bentuk model MARS dalam notasi matriks dituliskan pada persamaan (2.5) atau sebagai berikut:

$$Y = Ba + \varepsilon$$

$$\varepsilon = Y - Ba$$

dengan:

$$\mathbf{Y} = (y_1, y_2, \dots, y_n)'$$

$$\mathbf{a} = (a_1, a_2, \dots, a_M)'$$

$$\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)'$$

Pendugaan parameter dilakukan dengan meminimalkan jumlah kuadrat galat *Sum of Squared Error* (SSE) yang dirumuskan sebagai berikut:

$$\begin{aligned} SSE &= \sum_{i=1}^n \varepsilon_i^2 = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} \\ &= (\mathbf{Y} - \mathbf{B}\mathbf{a})^T (\mathbf{Y} - \mathbf{B}\mathbf{a}) \\ &= (\mathbf{Y}^T - \mathbf{a}^T \mathbf{B}^T) (\mathbf{Y} - \mathbf{B}\mathbf{a}) \\ &= \mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \mathbf{B}\mathbf{a} - \mathbf{a}^T \mathbf{B}^T \mathbf{Y} + \mathbf{a}^T \mathbf{B}^T \mathbf{B}\mathbf{a} \\ &= \mathbf{Y}^T \mathbf{Y} - 2\mathbf{a}^T \mathbf{B}^T \mathbf{Y} + \mathbf{a}^T \mathbf{B}^T \mathbf{B}\mathbf{a} \end{aligned}$$

Dalam memperoleh persamaan normal, fungsi SSE diturunkan secara parsial terhadap parameter a . Proses diferensiasi ini bertujuan untuk mendapatkan nilai parameter yang optimal dengan menyamakan hasil turunan tersebut dengan nol, sehingga diperoleh persamaan normal sebagai berikut:

$$\begin{aligned} \frac{\partial SSE}{\partial a} &= \frac{\partial}{\partial a} (\mathbf{Y}^T \mathbf{Y}) - \frac{\partial}{\partial a} (2\mathbf{a}^T \mathbf{B}^T \mathbf{Y}) + \frac{\partial}{\partial a} (\mathbf{a}^T \mathbf{B}^T \mathbf{B}\mathbf{a}) = 0 \\ 0 - 2\mathbf{B}^T \mathbf{Y} + 2\mathbf{B}^T \mathbf{B}\mathbf{a} &= 0 \\ -2\mathbf{B}^T \mathbf{Y} + 2\mathbf{B}^T \mathbf{B}\mathbf{a} &= 0 \\ -\mathbf{B}^T \mathbf{Y} + \mathbf{B}^T \mathbf{B}\mathbf{a} &= 0 \\ \mathbf{B}^T \mathbf{B}\mathbf{a} - \mathbf{B}^T \mathbf{Y} &= 0 \\ \mathbf{B}^T \mathbf{B}\mathbf{a} &= \mathbf{B}^T \mathbf{Y} \end{aligned}$$

Sehingga solusi pendugaan parameter dapat diperoleh sebagai berikut:

$$\mathbf{a} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{Y}$$

Dengan asumsi bahwa matriks $\mathbf{B}^T \mathbf{B}$ bersifat nonsingular, maka estimator parameter model MARS dapat dirumuskan sebagai berikut:

$$\hat{a} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T Y \quad (2.11)$$

2.4 Pemilihan Model MARS Terbaik

Pada tahap *backward pruning* akan dipilih satu kombinasi *basis function* dan mengeliminasi *basis function* yang memiliki kontribusi kecil terhadap model. Kemudian proses berlanjut hingga tidak ada *basis function* yang tereliminasi. Model MARS terbaik diperoleh berdasarkan nilai *Generalized Cross Validation* (GCV) dari model tersebut mempunyai nilai yang paling rendah (minimum) diantara model-model yang lain (Friedman, 1991). Fungsi GCV untuk memilih model terbaik yaitu dengan rumus sebagai berikut:

$$\text{GCV}(M) = \frac{\frac{1}{n} \sum_{i=1}^n [y_i - \hat{f}_M(x_i)]^2}{\left(1 - \frac{C(M)}{n}\right)^2} = \frac{\text{MSE}}{\left(1 - \frac{C(M)}{n}\right)^2} \quad (2.12)$$

dimana:

$$C(M) = K + d(K - 1)$$

dengan:

MSE = Mean Squared Error

y_i = variabel respon pada pengamatan ke- i .

x_i = variabel prediktor pada pengamatan ke- i .

n = banyaknya pengamatan.

M = banyaknya *basis function*.

$C(M)$ = ukuran kompleksitas model.

$\hat{f}_M(x_i)$ = nilai prediksi variabel respon.

K = banyaknya *basis function*.

d = faktor penalti kompleksitas model (bernilai antara 2 hingga 3).

Parameter d merupakan faktor penalti kompleksitas model yang digunakan untuk mengontrol tingkat kompleksitas model MARS. Nilai d umumnya berada pada rentang 2 hingga 3, di mana nilai yang lebih besar memberikan penalti yang lebih kuat terhadap model yang memiliki banyak fungsi basis sehingga dapat mengurangi risiko *overfitting* dan meningkatkan kemampuan generalisasi model.

2.5 Pendapatan Daerah

Pendapatan daerah merupakan seluruh penerimaan yang menjadi hak pemerintah daerah dan menambah kekayaan bersih dalam satu tahun anggaran, tanpa kewajiban untuk dikembalikan. Pendapatan ini dicatat melalui rekening kas umum daerah dan dianggarkan dalam APBD sebagai perkiraan yang terukur secara rasional. Pendapatan daerah mencerminkan kemampuan suatu daerah dalam membiayai kegiatan pemerintahan dan Pembangunan (Halim, 2014).

Menurut Undang-Undang No. 33 Tahun 2004, sumber pendapatan daerah terdiri atas tiga kelompok utama, yaitu:

1. Pendapatan Asli Daerah (PAD) merupakan penerimaan yang bersumber dari potensi ekonomi daerah sendiri seperti pajak, retribusi, dan hasil pengelolaan kekayaan daerah.
2. Transfer ke Daerah dan Dana Desa (TKDD) merupakan bagian dari Belanja Negara yang dialokasikan dalam anggaran pendapatan dan belanja negara kepada Daerah dan Desa dalam rangka mendanai pelaksanaan urusan yang telah diserahkan kepada daerah dan desa.
3. Pendapatan Lainnya merupakan penerimaan daerah di luar PAD dan pendapatan transfer dari pemerintah pusat, yang tetap sah menurut peraturan perundang-undangan. Komponen ini mencakup hibah, dana darurat, pendapatan lain sesuai ketentuan, dan transfer antar daerah.

2.6 Kepadatan Penduduk

Menurut Badan Pusat Statistik (BPS) Kepadatan penduduk adalah suatu keadaan yang dikatakan semakin padat bila jumlah manusia pada suatu batas ruang tertentu semakin banyak dibandingkan dengan luas ruangnya. Nilai kepadatan penduduk dinyatakan dalam satuan orang (jiwa) dan termasuk dalam ukuran rasio. Dalam penyajiannya, indikator ini dapat diklasifikasikan menjadi tiga kategori, yaitu kepadatan rendah (< 150 jiwa/ha), sedang (151–200 jiwa/ha), dan tinggi (201–400 jiwa/ha). Indikator ini bukan merupakan indikator komposit karena hanya dibangun

oleh dua variabel utama, yaitu jumlah penduduk dan luas lahan. Berikut merupakan metode perhitungan untuk kepadatan penduduk.

$$\text{Kepadatan Penduduk} = \frac{\text{Jumlah Penduduk}}{\text{Luas Lahan}}$$

Dalam konteks keuangan daerah, Kepadatan penduduk berpengaruh terhadap pendapatan daerah karena semakin padat suatu wilayah, umumnya semakin tinggi aktivitas ekonomi dan basis pajak yang meningkatkan Pendapatan Daerah. Namun, dampak ini bergantung pada kualitas pembangunan, karena kepadatan yang tidak terkelola dapat justru menekan potensi pendapatan daerah.

2.7 Produk Domestik Regional Bruto (PDRB) Per Kapita

Produk Domestik Regional Bruto (PDRB) per kapita merupakan rata-rata pendapatan yang diperoleh penduduk di suatu wilayah (provinsi/kabupaten/kota) dalam periode tertentu, yang dihitung dengan membagi total Produk Domestik Regional Bruto (PDRB) dengan jumlah penduduk. PDRB sendiri adalah nilai tambah bruto seluruh barang dan jasa yang dihasilkan suatu wilayah yang timbul akibat berbagai aktivitas ekonomi dalam suatu periode tertentu. Penyusunan PDRB dapat dilakukan melalui tiga pendekatan yaitu pendekatan produksi, pengeluaran, dan pendapatan yang disajikan atas dasar harga berlaku dan harga konstan (riil) yang digunakan untuk melihat struktur ekonomi dan pertumbuhan ekonomi daerah (BPS, 2024).

2.8 Indeks Pembangunan Manusia (IPM)

Indeks Pembangunan Manusia (IPM) merupakan pengukuran perbandingan dari harapan hidup (kesehatan), pendidikan, dan standar hidup untuk semua negara. IPM digunakan sebagai indikator untuk menilai aspek kualitas dari pembangunan dan untuk mengklasifikasikan apakah sebuah negara termasuk negara maju, negara berkembang, atau negara terbelakang dan juga untuk mengukur pengaruh dari kebijakan ekonomi terhadap kualitas hidup (BPS, 2015).

Dalam konteks ekonomi daerah, IPM berperan penting karena kualitas sumber daya manusia memengaruhi produktivitas tenaga kerja dan aktivitas ekonomi. Daerah dengan IPM tinggi cenderung memiliki tingkat pendapatan masyarakat yang lebih besar, yang pada akhirnya dapat meningkatkan pendapatan daerah melalui peningkatan konsumsi, pajak, dan aktivitas ekonomi lokal (Lestari dkk., 2024)

2.9 Penduduk Miskin

Secara umum, kemiskinan merupakan kondisi dimana seseorang atau sekelompok orang tidak mampu memenuhi hak-hak dasarnya untuk mempertahankan dan mengembangkan kehidupan yang bermartabat (Bappenas, 2018). Penduduk miskin dapat diartikan kondisi ketidakmampuan masyarakat dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan, yang memiliki rata-rata pengeluaran per kapita perbulan di bawah Garis Kemiskinan (GK). Data kemiskinan yang dirilis BPS merupakan hasil survei sosial ekonomi nasional (BPS, 2023)

Menurut Kuncoro (2010), kemiskinan memiliki hubungan erat dengan kapasitas fiskal daerah. Daerah dengan tingkat kemiskinan tinggi umumnya memiliki aktivitas ekonomi rendah, yang berdampak pada kecilnya kontribusi terhadap Pendapatan Daerah. Sebaliknya, penurunan angka kemiskinan akan meningkatkan partisipasi ekonomi masyarakat, sehingga memperluas basis pajak dan meningkatkan penerimaan daerah.

BAB III

METODE PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester ganjil tahun akademik 2025/2026 bertempat di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder, yaitu data Pendapatan Daerah (Y), Kepadatan Penduduk (X_1), Produk Domestik Regional Bruto (PDRB) Per Kapita (X_2), Indeks Pembangunan Manusia (X_3), dan Penduduk Miskin (X_4) per kabupaten/kota di Provinsi Lampung pada tahun 2015-2024. Data yang digunakan diperoleh dari publikasi Portal Data APBD Direktorat Jenderal Perimbangan Keuangan (DJPK) Kementerian Keuangan serta publikasi Badan Pusat Statistik (BPS). Data penelitian dapat diakses melalui link berikut: https://docs.google.com/spreadsheets/u/0/d/114Iu1fHtC1V4zSBgCXT_CjnHEH7SSrcV/htmlview

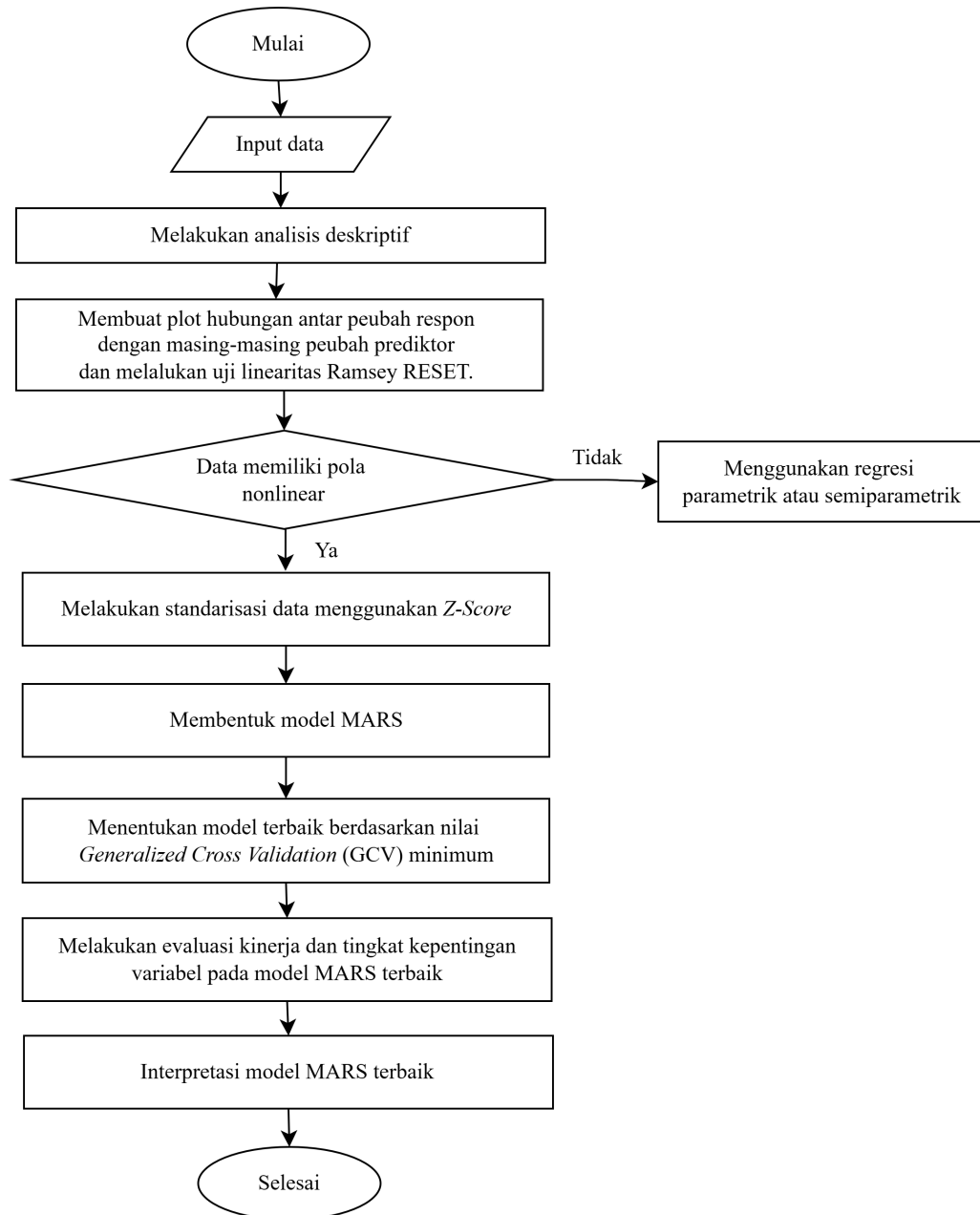
3.3 Metode Penelitian

Penelitian ini menerapkan metode *Multivariate Adaptive Regression Splines* (MARS) dalam memodelkan pendapatan daerah Provinsi Lampung. Proses penelitian dilakukan melalui beberapa tahapan yang sistematis, yaitu mulai dari pengumpulan data, pengolahan data, analisis deskriptif hingga pembentukan model dengan bantuan *software R-Studio*.

Adapun langkah-langkah yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Melakukan analisis deskriptif terhadap variabel penelitian. Hal ini bertujuan untuk mengetahui bagaimana karakteristik dari data pendapatan daerah di Provinsi Lampung.
2. Membuat plot antara variabel pendapatan daerah (Y) dengan variabel Kepadatan Penduduk (X_1), PDRB per kapita (X_2), IPM (X_3), dan penduduk miskin (X_4) yang dijadikan sebagai deteksi awal mengenai pola hubungan antara variabel dependen dengan variabel independen. Selain itu, dilakukan uji linearitas menggunakan uji *Ramsey RESET (Regression Specification Error Test)*. Jika hasil menunjukkan nonlinear atau hubungan yang tidak teratur, maka analisis dilanjutkan dengan metode MARS.
3. Dalam penelitian ini sebelum dilakukan pemodelan, data terlebih dahulu distandarisasi menggunakan metode *Z-score*. Standarisasi ini bertujuan untuk menyamakan skala antar variabel sehingga proses pemodelan menjadi lebih stabil dan tidak dipengaruhi oleh perbedaan satuan data.
4. Membentuk model dengan data yang sudah distandarisasi. Pemodelan dilakukan dengan menentukan maksimum *Basis function* (BF). Friedman (1990) menyarankan maksimum *basis function* yang digunakan antara 2-4 kali ukuran variabel independen. Variabel independen yang digunakan dalam penelitian ini sebanyak 4 variabel maka maksimum *basis function* yang dibentuk adalah 8 sampai 16.
5. Menentukan jumlah Maksimum Interaksi (MI) yaitu 1, 2, dan 3 interaksi.
6. Menentukan jumlah Minimum Observasi (MO) yaitu 0, 1, 2 dan 3 observasi.
7. Menentukan model terbaik dengan memperhatikan nilai *Generalized Cross Validation* (GCV). Nilai GCV yang kecil, maka semakin baik model tersebut dibandingkan model lainnya.
8. Melakukan evaluasi kinerja dan tingkat kepentingan variabel pada model MARS terbaik.
9. Melakukan interpretasi hasil analisis.

Berikut adalah diagram alir tahapan penelitian yang dilakukan dalam memodelkan pendapatan daerah Provinsi Lampung menggunakan metode *Multivariate Adaptive Regression Splines* (MARS).



Gambar 3.1 Diagram Alir Penelitian

BAB V

KESIMPULAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan mengenai pemodelan pendapatan daerah menggunakan metode MARS pada data kabupaten/kota di Provinsi Lampung diperoleh kesimpulan sebagai berikut:

1. Model MARS terbaik dengan nilai GCV terkecil yaitu 0,0623 dengan kombinasi nilai BF = 12, MI = 3 dan MO = 1, dengan model yang melibatkan variabel kepadatan penduduk, PDRB per kapita, Indeks Pembangunan Manusia (IPM), dan penduduk miskin dalam menjelaskan pendapatan daerah di Provinsi Lampung.
2. Variabel yang paling berpengaruh dalam pembentukan model MARS terbaik adalah penduduk miskin dengan nilai kepentingan sebesar 100%. Selanjutnya diikuti oleh variabel PDRB per kapita sebesar 45,4%, IPM sebesar 27,3%, dan variabel kepadatan penduduk sebesar 22,3%.

DAFTAR PUSTAKA

- Budiantara, I. N. 2005. Model keluarga spline polinomial truncated dalam regresi semiparametrik. *Berkala Mipa*. **15**(3): 55-61.
- Badan Pusat Statistik. Kepadatan Penduduk. Sistem Informasi Rujukan Statistik (Sirusa). Diakses disini.
- Badan Pusat Statistik. 2024. Produk Domestik Regional Bruto. Sistem Informasi Rujukan Statistik (Sirusa). Diakses disini.
- Badan Pusat Statistik. 2015. Indeks Pembangunan Manusia. Sistem Informasi Rujukan Statistik (Sirusa). Diakses disini.
- Badan Pusat Statistik. 2023. Penghitungan dan Analisis Kemiskinan Makro Indonesia. Sistem Informasi Rujukan Statistik (Sirusa). Diakses disini.
- Badan Perencanaan Pembangunan Nasional. 2018. Profil Kemiskinan di Indonesia. Jakarta: Bappenas.
- Eubank, R. L. 1999. *Nonparametrik regression and spline smoothing*. Marcell Dekker, New York.
- Friedman, J. H. 1991. Multivariat adaptive regression splines. *The Annals of Statistics*. **19**(1): 1-67.
- Gujarati, Damodar N., dan Dawn C.P. 2009. *Basic Econometrics*. The McGraw-Hill Companies Inc, Singapura.
- Halim, MS Kusufi. 2014. *Akuntansi Sektor Publik*. Salemba Empat, Jakarta.
- Hardle, W. 1992. *Applied Nonparametric Regression*. Cambridge University Press, New York.
- Hasan, M. I. 2001. *Pokok pokok Materi Statistika 1 (Statistika Deskriptif)*. Bumi Aksara, Jakarta.

- Kartini, A. Y. K., & Ummah, L. N. 2022. Pemodelan Kejadian Balita Stunting di Kabupaten Bojonegoro dengan Metode Geographically Weighted Regression (GWR) dan Multivariate Adaptive Regression Splines (MARS). *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*. **15**(1): 133-134.
- Kuncoro, M. 2010. *Otonomi dan pembangunan daerah: reformasi, perencanaan, strategi dan peluang*. Erlangga, Jakarta.
- Kuswanto, D., 2012. *Statistik untuk pemula & orang awam*. Laskar Aksara, Jakarta.
- Lestari, A., Frinaldi, A., & Wahyuni, Y. S. 2024. Optimalisasi Pendapatan Asli Daerah (PAD) Melalui Inovasi Pendidikan. *Menara Ilmu: Jurnal Penelitian dan Kajian Ilmiah*. **18**(1): 12-22.
- Mardiasmo, M.B.A. 2021. *Otonomi & manajemen keuangan daerah* . edisi terbaru. Penerbit Andi, Daerah Istimewa Yogyakarta.
- Mattalunru, M. R., Annas, S., & Aidid, M. K. 2022. Aplikasi Multivariate Adaptive Regression Splines (MARS) Untuk Mengetahui Faktor yang Mempengaruhi Curah Hujan di Kota Makassar. *VARIANSI: Journal of Statistics and Its application on Teaching and Research*. **4**(1): 9-19.
- Nash, M. S., & Bradford, D. F. 2001. *Parametric and nonparametric logistic regressions for prediction of presence/absence of an amphibian*. US Environmental Protection Agency.
- Otok, B. W. 2010. Pendekatan Multivariate Adaptive Regression Spline (MARS) Pada Pengelompokan Zona Musim Suatu Wilayah, *Statistika*. **10**(2): 107-120.
- Panggabean, G. F., & Mansyur, A. 2023. Penerapan Multivariate Adaptive Regression Splines pada Laju Produk Domestik Regional Bruto Menurut Lapangan Usaha di Provinsi Sumatera Utara. *Jurnal Sains Dan Teknologi*. **2**(1), 159-171.
- Pratiwi, P. I. 2020. *Perbandingan Regresi Logistik Biner dan Multivariate Adaptive Regression Spline (MARS) Terhadap Faktor-Faktor Yang Mempengaruhi Tingkat Kelahiran Di Kota Kendari*. Disertasi, Muhammadiyah University, Semarang.
- Ramsey, J. B. 1969. *Tests for Specification Errors in Classical Linear Least Squares Regression Analysis*. *Journal of the Royal Statistical Society: Series B (Methodological)*, **31**(2), 350–371.

- Sita, E. D. A. A & Otok, B. W. 2014. Pendekatan Multivariate Adaptive Regression Spline (MARS) pada Pemodelan Penduduk Miskin di Indonesia Tahun 2008-2012. *Jurnal Seminar Nasional Matematika*.
- Sjafrizal. 2008. *Ekonomi Regional Teori dan Aplikasi*. Baduose Media. Padang, Sumatera Barat.
- Sugiyono, 2007. *Metode Penelitian Kualitatif Kuantitatif dan R&D*. Alfabeta, Bandung.
- Undang-Undang No. 33 Tahun 2004 mengatur tentang Perimbangan Keuangan antara Pemerintah Pusat dan Pemerintahan Daerah.
- Walpole, R. E. 1995. *Pengantar statistika*. Gramedia, Jakarta.
- Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. 2012. *Probability and Statistics for Engineers and Scientists*. 9th Edition. Pearson Education.
- Zakiyah Mar'ah, Ruliana, and Magfirah Septiana. 2024. Pemodelan Multivariate Adaptive Regression Spline (MARS) Pada Indeks Harga Saham Gabungan (IHSG) Tahun 2018–2023. *Journal of Statistics and Its application on Teaching and Research* **6**(1): 1-10.