

**PEMODELAN *MULTIVARIATE ADAPTIVE REGRESSION***

***SPLINE (MARS)***

(Studi Kasus: Kasus Diare Pada Balita di Provinsi Jawa Barat-Jawa Tengah  
Tahun 2019)

**SKRIPSI**

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



diajukan oleh:

A'yun Nafsi Utami

16610020

Kepada

**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UIN SUNAN KALIJAGA  
YOGYAKARTA  
2021**

# SURAT PERSETUJUAN



Universitas Islam Negeri Sunan Kalijaga



FM-UINSK-BM-05-03/R0

## SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi / Tugas Akhir

Lamp :

Kepada

Yth. Dekan Fakultas Sains dan Teknologi

UIN Sunan Kalijaga Yogyakarta

di Yogyakarta

*Assalamu'alaikum wr. wb.*

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : A'yun Nafsi Utami  
NIM : 16610020  
Judul Skripsi : Pemodelan *Multivariate Adaptive Regression Spline* (MARS)

sudah dapat diajukan kembali kepada Program Studi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Matematika.

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

*Wassalamu'alaikum wr. wb.*

Yogyakarta, 19 Maret 2021

Mengetahui,

Pembimbing

M. Farhan Qudratullah, S. Si., M. Sc

NIP. 19790922 200801 1 011

# PENGESAHAN TUGAS AKHIR



KEMENTERIAN AGAMA  
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA  
FAKULTAS SAINS DAN TEKNOLOGI  
Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

## PENGESAHAN TUGAS AKHIR

Nomor : B-857/Un.02/DST/PP.00.9/05/2021

Tugas Akhir dengan judul : PEMODELAN MULTIVARIATE ADAPTIVE REGRESSION SPLINE (MARS) (Studi Kasus : Diare pada Balita di Provinsi Jawa Barat - Jawa Tengah Tahun 2019)

yang dipersiapkan dan disusun oleh:

Nama : A'YUN NAFSI UTAMI  
Nomor Induk Mahasiswa : 16610020  
Telah diujikan pada : Rabu, 07 April 2021  
Nilai ujian Tugas Akhir : A-

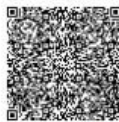
dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

## TIM UJIAN TUGAS AKHIR



Ketua Sidang  
Mohammad Farhan Qudratullah, S.Si., M.Si  
SIGNED

Valid ID: 60ac6b65e72e



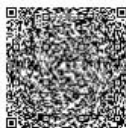
Penguji I  
Sri Utami Zuliana, S.Si., M.Sc., Ph.D.  
SIGNED

Valid ID: 60876e24dc28



Penguji II  
Dr. Epha Diana Supandi, S.Si., M.Sc.  
SIGNED

Valid ID: 6087a7461fa57



Yogyakarta, 07 April 2021  
UIN Sunan Kalijaga  
Dekan Fakultas Sains dan Teknologi  
Dr. Dra. Hj. Khurul Wardati, M.Si.  
SIGNED

Valid ID: 60168594e4f47

## SURAT PERNYATAAN KEASLIAN

### SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : A'yun Nafsi Utami

NIM : 16610020

Program Studi : Matematika

Fakultas : Sains dan Teknologi

Dengan ini menyatakan bahwa isi skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu Perguruan Tinggi dan sesungguhnya skripsi ini merupakan hasil pekerjaan penulis sendiri sepanjang pengetahuan penulis, bukan duplikasi dari karya orang lain kecuali bagian tertentu yang penulis ambil sebagai bahan acuan. Apabila terbukti pernyataan ini tidak benar, sepenuhnya menjadi tanggung jawab penulis.

Yogyakarta, 25 Maret 2021

Yang Menyatakan  
  
A'yun Nafsi Utami



**MOTTO**

*“You Can Do It If You Believe”*

## **PERSEMBAHAN**

Skripsi ini saya persembahkan untuk:

Kedua Orang Tua Saya

Kedua Adik Saya serta Keluarga Besar

dan

Almamater tercinta, Universitas Islam Negeri Sunan Kalijaga Yogyakarta

## KATA PENGANTAR

Puji syukur penulis haturkan kepada Allah SWT yang telah melimpahkan nikmat dan hidayah-Nya sehingga penulis bisa menyelesaikan penelitian tugas akhir ini yang berjudul “Pemodelan *Multivariate Adaptive Regression Spline* (MARS)”. Shalawat serta salam penulis haturkan kepada Nabi Muhammad SAW junjungan kita beserta keluarga dan para pengikutnya.

Penelitian tugas akhir ini disusun sebagai salah satu syarat untuk memenuhi kelulusan mahasiswa Strata Satu Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta. Tujuan penulisan penelitian ini untuk memberikan wawasan mengenai penerapan ilmu statistika, salah satunya *Multivariate Adaptive Regression Spline* (MARS).

Penyusunan penelitian tugas akhir ini terwujud karena adanya dukungan, bimbingan, dan motivasi dari berbagai pihak baik secara moril dan materiil. Oleh karena itu, pada kesempatan kali ini penulis ingin menyampaikan ucapan terimakasih yang sebesar-besarnya kepada :

1. Bapak Prof. Dr. Phil. Al Makin, S. Ag., M. A., selaku Rektor Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
2. Ibu Dr. Khurul Wardati, M. Si., selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
3. Bapak Muchammad Abrori, S. Si., M. Kom., selaku ketua Program Studi Matematika Universitas Islam Negeri Sunan Kalijaga Yogyakarta.

4. Ibu Dr. Epha Diana Supandi, M Si., selaku pembimbing akademik yang telah memberikan pelayanan dan kelancaran akademik.
5. Bapak Mohammad Farhan Quadratullah, S. Si., M. Si., selaku pembimbing skripsi yang telah memberikan ilmu pengetahuan dan pengalaman serta waktu yang berharga kepada penulis sehingga dapat memudahkan dalam penelitian skripsi ini.
6. Bapak/ Ibu Dosen dan Staff Fakultas Sains dan Teknologi UIN Sunan Kalijaga atas ilmu dan bimbingannya selama ini, serta pelayanan selama perkuliahan sehingga selesainya skripsi ini.
7. Kedua orang tua penulis Panut Ngadiyono dan Kamini yang selalu memberikan semangat, kasih sayang dan do'a yang selalu mengiringi di setiap langkah penulis, pengorbanan yang telah diberikan kepada penulis selama ini.
8. Adik-adik penulis (Juta Hasby Wijaya dan Denisa Naffa Kurbatan) serta semua saudara yang telah memberikan semangat, do'a, dan motivasi.
9. Mas Rofiq Maulana Sya'ban yang selalu memberikan semangat, do'a, dan bantuan dalam menyelesaikan skripsi ini.
10. Teman-teman dari kecil, Detta Ramadhani, Fatlun Kusuma, dan Ilma Akhida yang selalu memberikan motivasi dalam menyelesaikan skripsi ini kepada penulis dengan cara kalian.
11. Teman-teman dari SMP, Ayuk dan Rahma yang selalu memberikan motivasi, dukungan, dan do'a untuk penulis.



12. Teman-teman kuliah, Saskia Ayu Gunawan dan Astika Riawa Putri yang selalu memberikan semangat saling mendorong agar penulis cepat menyelesaikan tugas akhir.
13. Teman-teman Matematika angkatan 2016, terimakasih sudah berjuang bersama selama 4 tahun ini untuk mewujudkan cita-cita.
14. Teman-teman KKN 185 Dusun Tompak, Giritirto, Purwosari, Gunungkidul yang sudah memberikan pengalaman menarik selama masa-masa KKN.
15. Semua pihak yang tidak dapat penulis sebutkan satu per satu yang telah membantu dalam menyelesaikan tugas akhir ini.

Penulis menyadari bahwa tugas akhir ini masih jauh dari kata sempurna. Oleh karena itu, penulis memohon maaf apabila terdapat kekurangan dalam penulisan skripsi ini. Penulis menerima kritik dan saran yang membangun demi perbaikan tugas akhir ini. Penulis berharap penelitian ini dapat bermanfaat dan memberikan pengetahuan bagi penulis dan semua pihak yang membutuhkan.

Yogyakarta, 29 Maret 2021

**A'yun Nafsi Utami**

## DAFTAR ISI

<b>COVER</b> .....	<b>i</b>
<b>SURAT PERSETUJUAN TUGAS AKHIR</b> .....	<b>ii</b>
<b>PENGESAHAN TUGAS AKHIR</b> .....	<b>iii</b>
<b>SURAT PERNYATAAN KEASLIAN</b> .....	<b>iv</b>
<b>MOTTO</b> .....	<b>v</b>
<b>PERSEMBAHAN</b> .....	<b>vi</b>
<b>KATA PENGANTAR</b> .....	<b>vii</b>
<b>DAFTAR ISI</b> .....	<b>x</b>
<b>DAFTAR SIMBOL</b> .....	<b>xv</b>
<b>DAFTAR GAMBAR</b> .....	<b>xvii</b>
<b>DAFTAR TABEL</b> .....	<b>xviii</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xix</b>
<b>INTISARI</b> .....	<b>xx</b>
<b>ABSTRACT</b> .....	<b>xxi</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1    Latar Belakang Masalah .....	1
1.2    Rumusan Masalah .....	6
1.3    Tujuan Penelitian.....	6
1.4    Batasan Masalah.....	7
1.5    Manfaat Penelitian.....	7
1.5.1    Bagi Mahasiswa .....	7
1.5.2    Bagi Pembaca.....	8

1.6	Tinjauan Pustaka .....	8
1.7	Sistematika Penulisan.....	12
<b>BAB II LANDASAN TEORI .....</b>		<b>14</b>
2.1	Matriks.....	14
2.1.1	Operasi Matriks .....	14
2.1.2	Transpose Matriks.....	16
2.1.3	Invers Matriks .....	17
2.1.4	Trace Matriks .....	17
2.2	Teori Probabilitas .....	18
2.3	Variabel Random.....	21
2.4	Fungsi Distribusi Probabilitas .....	21
2.5	Karakteristik Distribusi Probabilitas .....	24
2.5.1	Ekspetasi .....	24
2.5.2	Variansi .....	24
2.5.3	Sifat-sifat Mean dan Variansi.....	25
2.6	Distribusi Normal .....	25
2.7	Estimasi <i>Maximum Likelihood</i> .....	26
2.8	Model Analisis Regresi .....	28
2.8.1	Asumsi Model Regresi Linear .....	32
2.8.2	Estimasi Regresi Linear Sederhana dengan <i>Maximum Likelihood</i> . 33	
2.8.3	Sifat-sifat Estimator Regresi Linear Sederhana .....	35
2.8.4	Pengujian Parameter Model .....	41
2.8.5	Penyimpangan dalam Analisis Regresi .....	44

2.8.6	Asumsi Model Regresi.....	45
2.9	Model Regresi Linear Berganda.....	49
2.10	Regresi Nonparametrik.....	50
2.11	Regresi <i>Spline</i> .....	53
2.12	Diare .....	55
2.12.1	Klasifikasi Diare.....	55
2.12.2	Faktor Yang Mempengaruhi Terjadinya Diare .....	56
<b>BAB III METODE PENELITIAN .....</b>		<b>59</b>
3.1	Studi Pustaka .....	59
3.2	Jenis dan Sumber Data .....	59
3.3	Variabel .....	60
3.4	Tahapan Penelitian .....	60
3.5	<i>Flowchart</i> Penelitian .....	62
<b>BAB IV PEMBAHASAN.....</b>		<b>63</b>
4.1	Model Regresi Nonparametrik .....	63
4.2	Regresi <i>Spline</i> .....	64
4.3	Model <i>Multivariate Adaptive Regression Spline</i> (MARS).....	66
4.4	Estimasi Parameter Menggunakan <i>Maximum Likelihood</i> .....	72
4.5	Pemilihan Model dengan <i>Generalized Cross Validation</i> (GCV).....	79
4.6	Algoritma MARS .....	80
4.7	Pemilihan Model MARS Terbaik.....	82
4.8	Pengujian Signifikansi Model MARS .....	85
<b>BAB V STUDI KASUS .....</b>		<b>87</b>

5.1	Deskriptif Data .....	87
5.2	<i>Scatterplot</i> Jumlah Kasus Penyakit Diare dengan Faktor-Faktor yang Diduga Mempengaruhi.....	90
5.3	Uji Parameter Kelayakan Model .....	93
5.3.1	Pengujian Serentak (Uji F).....	93
5.3.2	Pengujian Individu (Uji t) .....	94
5.4	Uji Asumsi Model Regresi .....	96
5.4.1	Uji Normalitas .....	96
5.4.2	Uji Heterokedastisitas .....	97
5.4.3	Uji Autokorelasi .....	98
5.4.4	Uji Multikolinearitas .....	100
5.5	Pembentukan Model Regresi.....	100
5.6	Model MARS Terbaik.....	101
5.7	Interpretasi Model MARS .....	102
5.8	Tingkat Kepentingan Variabel Prediktor .....	103
5.9	Pengujian Signifikansi Model Terbaik MARS .....	105
5.9.1	Uji Simultan (Uji Serentak) .....	105
5.9.2	Uji Parsial.....	106
5.10	Pembahasan .....	108
<b>BAB VI KESIMPULAN DAN SARAN.....</b>		<b>110</b>
6.1	Kesimpulan.....	110
6.2	Saran .....	111
<b>DAFTAR PUSTAKA .....</b>		<b>113</b>

<b>LAMPIRAN.....</b>	<b>117</b>
<b>DAFTAR RIWAYAT HIDUP .....</b>	<b>128</b>

## DAFTAR SIMBOL

- $Y$  : data jumlah kasus diare di Jawa Tengah-Jawa Barat
- $X_1$  : data Tempat Pengelolaan Makanan (TPM) sehat
- $X_2$  : data air minum yang memenuhi syarat kesehatan
- $X_3$  : data pemberian ASI Eksklusif pada bayi usia kurang dari 6 bulan
- $ZY$  : data jumlah kasus diare di Jawa Tengah-Jawa Barat setelah di standarisasi
- $ZX_1$  : data Tempat Pengelolaan Makanan (TPM) sehat setelah di standarisasi
- $ZX_2$  : data air minum yang memenuhi syarat kesehatan setelah di standarisasi
- $ZX_3$  : data pemberian ASI Eksklusif pada bayi usia kurang dari 6 bulan setelah di standarisasi
- $f(x_i)$  : fungsi smooth yang tidak diketahui
- $BF$  : fungsi basis
- $MI$  : maksimum interaksi
- $MO$  : minimum observasi
- $n$  : jumlah sampel
- $M$  : maksimum fungsi basis ke- $m$
- $y_i$  : variabel respon pada amatan ke- $i$
- $\beta_0, \beta_1$  : parameter koefisien regresi
- $a_0$  : fungsi basis induk

- $a_m$  : koefisien dari fungsi basis ke- $m$
- $k_m$  : derajat interaksi ke- $m$
- $S_{km}$  : nilai 1 atau -1 jika data berada di sebelah kanan atau kiri knot
- $z_{km}$  : nilai knots dari variabel prediktor  $z_{p_n(k,m)}$
- $z_{p_n(k,m)}$  : variabel prediktor dari  $p$  dengan observasi ke- $n$
- $\varepsilon_i$  : error ke-  $i$



## DAFTAR GAMBAR

<b>Gambar 3. 1</b> <i>Flowchart</i> Penelitian.....	<b>62</b>
<b>Gambar 4. 1</b> Fungsi <i>Spline</i> Linear dengan satu titik knot .....	<b>66</b>
<b>Gambar 5.1</b> <i>Boxplot</i> Variabel Jumlah Kasus Kejadian Diare pada Balita .....	<b>88</b>
<b>Gambar 5.2</b> <i>Boxplot</i> Variabel Tempat Pengelolaan Makanan (TPM) yang Memenuhi Syarat Kesehatan .....	<b>89</b>
<b>Gambar 5.3</b> <i>Boxplot</i> Variabel Jumlah Air Minum yang Memenuhi Syarat Kesehatan .....	<b>89</b>
<b>Gambar 5.4</b> <i>Boxplot</i> Variabel Pemberian ASI Eksklusif pada Bayi Usia Kurang dari 6 Bulan .....	<b>90</b>
<b>Gambar 5.5</b> <i>Scatterplot</i> Antara Jumlah Kasus Penyakit Diare Pada Balita dengan Tempat Pengelolaan Makanan (TPM) yang Memenuhi Syarat Kesehatan .....	<b>91</b>
<b>Gambar 5.6</b> <i>Scatterplot</i> Antara Jumlah Kasus Diare Pada Balita dengan Akses Air Minum yang Memenuhi Syarat Kesehatan .....	<b>92</b>
<b>Gambar 5.7</b> <i>Scatterplot</i> Antara Jumlah Kasus Diare Dengan Pemberian ASI Eksklusif pada Bayi < 6 bulan .....	<b>93</b>
<b>Gambar 5.8</b> <i>Normal P-P Plot of Regression Standardized Residual</i> .....	<b>96</b>

## DAFTAR TABEL

Tabel 1.1 Tinjauan Pustaka.....	10
Tabel 1.2 Tinjauan Pustaka.....	11
Tabel 3.1 Variabel Respon dan Prediktor Penelitian .....	60
Tabel 5.1 Statistik Deskriptif Variabel.....	87
Tabel 5.2 Pengujian Serentak dengan Uji F.....	94
Tabel 5.3 Uji Koefisien Regresi.....	94
Tabel 5.4 <i>One Sample Kolmogorov Smirnov</i> .....	97
Tabel 5.5 Uji Heterokedastisitas .....	98
Tabel 5.6 <i>Durbin-Watson Test</i> .....	99
Tabel 5.7 Uji Multikolinieritas.....	100
Tabel 5.8 Tingkat Kepentingan Variabel.....	103
Tabel 5.9 Hasil Pengujian Secara Parsial.....	107

## DAFTAR LAMPIRAN

<b>Lampiran 1.</b> Tabel Data Penelitian.....	<b>117</b>
<b>Lampiran 2.</b> Tabel Data Penelitian Setelah Standarisasi.....	<b>120</b>
<b>Lampiran 3.</b> Script dan Output <i>Software R Boxplot</i> .....	<b>123</b>
<b>Lampiran 4.</b> Output SPSS Hasil Uji Simultan (Serentak).....	<b>124</b>
<b>Lampiran 5.</b> Output SPSS Hasil Uji Parsial .....	<b>124</b>
<b>Lampiran 6.</b> Script dan Output <i>Software R Uji Normalitas</i> .....	<b>124</b>
<b>Lampiran 7.</b> Script dan Output <i>Software R Uji Heterokedastisitas</i> .....	<b>125</b>
<b>Lampiran 8.</b> Script dan Output <i>Software R Uji Autokorelasi</i> .....	<b>125</b>
<b>Lampiran 9.</b> Script dan Output <i>Software R Uji Multikolinearitas</i> .....	<b>125</b>
<b>Lampiran 10.</b> Script dan Output <i>Software R Model MARS</i> .....	<b>126</b>
<b>Lampiran 11.</b> Script dan Output <i>Software R Tingkat Kepentingan</i> .....	<b>126</b>
<b>Lampiran 12.</b> Script dan Output <i>Software R Uji Simultan</i> .....	<b>127</b>
<b>Lampiran 13.</b> Script dan Output <i>Software R Uji Parsial</i> .....	<b>127</b>

**PEMODELAN *MULTIVARIATE ADAPTIVE REGRESSION*  
*SPLINE* (MARS)**

**(Studi kasus: Jumlah Kasus Diare Pada Balita di Provinsi Jawa Tengah-  
Jawa Barat Tahun 2019)**

**A'yun Nafsi Utami**

Jurusan Matematika Fakultas Sains dan Teknologi  
Universitas Islam Negeri Sunan Kalijaga

**INTISARI**

Regresi nonparametrik adalah salah satu metode statistika yang digunakan untuk mengetahui pola hubungan antara variabel respon dan variabel prediktor yang tidak diketahui bentuk fungsinya dan pola menyebar maka kurva regresi dapat diduga. *Multivariate Adaptive Regression Spline* (MARS) adalah salah satu metode analisis regresi nonparametrik yang digunakan untuk mengatasi permasalahan data yang berdimensi tinggi yaitu data yang memiliki jumlah variabel prediktor sebesar  $3 \leq p \leq 20$  dan data sampel yang berukuran  $50 \leq n \leq 1000$ .

Pada penelitian ini, penulis akan meneliti mengenai pemodelan *Multivariate Adaptive Regression Spline* (MARS) pada kasus diare di Provinsi Jawa Tengah-Jawa Barat tahun 2019. Untuk memilih model MARS terbaik, pada penelitian ini menggunakan nilai *Generalized Cross Validation* (GCV) terkecil atau minimum yang dihasilkan tiap model kombinasi dari minimum observasi (MO), *basis function* (BF) dan Interaksi (MI). Hasil analisis MARS pada kasus jumlah kasus diare pada balita di Jawa Tengah-Jawa Barat menghasilkan model terbaik dengan bentuk

$$\hat{Y} = -0,6236969 + 1,011132 * BF_1 - 1,308728 * BF_2 + 0,2991534 * BF_3 + 0,5829459 * BF_4$$

**Kata kunci:** regresi nonparametrik, MARS, GCV, kasus diare.

# MULTIVARIATE ADAPTIVE REGRESSION SPLINE (MARS) MODELING

(Case study: Number of Diarrhea Cases in Toddlers in Central Java-West  
Jawa Province in 2019)

A'yun Nafsi Utami

Department of Mathematics, Faculty of Science and Technology  
Sunan Kalijaga State Islamic University

## ABSTRACT

Nonparametric regression is one of the statistical methods used to determine the pattern of the relationship between the response variable and the predictor variable whose function is unknown and the spread pattern, so the regression curve can be predicted. Multivariate Adaptive Regression Spline (MARS) is a nonparametric regression analysis method that is used to solve high-dimensional data problems, namely data that has a number of predictor variables of  $3 \leq p \leq 20$  and sample data measuring  $50 \leq n \leq 1000$ .

In this study, the authors will examine the Multivariate Adaptive Regression Spline (MARS) modeling in cases of diarrhea in Central Java-West Java Province in 2019. To choose the best MARS model, this study uses the smallest or minimum Generalized Cross Validation (GCV) value. produced each model combination of minimum observation (MO), basis function (BF) and interaction (MI). The results of the MARS analysis on the number of cases of diarrhea in children under five in Central Java-West Java produce the best model with the form of the equation:  
$$\hat{Y} = -0,6236969 + 1,011132 * BF_1 - 1,308728 * BF_2 + 0,2991534 * BF_3 + 0,5829459 * BF_4$$

**Key words** : nonparametrics regression, MARS, GCV, cases of diarrhea.

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang Masalah**

Statistika merupakan salah satu cabang dari matematika terapan. Statistika adalah suatu disiplin ilmu yang mempelajari sekumpulan konsep dan metode pengumpulan, penyajian, analisis, dan interpretasi data sampai pada pengambilan keputusan. Statistika dalam pengertian yang paling sederhana adalah data. (Quadratullah, dkk, 2012).

Analisis regresi merupakan salah satu teknik analisis statistika yang sering digunakan untuk menyelesaikan permasalahan dalam statistika yang dapat menggambarkan ketergantungan atau mencari hubungan fungsional antara satu variabel respon (dependen) dengan variabel prediktor (independen). Variabel respon (dependen) adalah variabel terikat atau variabel yang dijelaskan oleh variabel lainnya. Sedangkan variabel prediktor (independen) adalah variabel bebas yang merupakan variabel penjelas, variabel yang memengaruhi atau variabel prediksi bagi variabel respon (dependen). Pendekatan dalam metode analisis regresi untuk mengestimasi kurva regresi dikelompokkan menjadi 3 (tiga), yaitu analisis regresi parametrik, analisis regresi nonparametrik, dan analisis regresi semi parametrik yang merupakan perpaduan dari regresi parametrik dan regresi nonparametrik. Dalam menjelaskan pola hubungan antara variabel respon dengan variabel prediktor, digunakan kurva regresi dengan pendekatan regresi parametrik, dimana diasumsikan bentuk kurva regresi diketahui (seperti linier, kuadratik, kubik) berdasarkan teori yang dapat memberikan informasi hubungan (Drapper, 1992).

Namun tidak semua pola hubungan dapat didekati dengan pendekatan parametrik, karena tidak adanya suatu informasi mengenai bentuk hubungan variabel respon dan variabel prediktor. Jika asumsi model parametrik tidak terpenuhi maka kurva regresi dapat diduga dengan menggunakan pendekatan model regresi nonparametrik, karena memiliki fleksibilitas yang tinggi dalam mengestimasi kurva regresi. Dalam pandangan regresi nonparametrik data diharapkan mencari sendiri estimasi kurva regresi tanpa dipengaruhi oleh faktor subyektifitas dari perancang penelitian (Eubank, 1999).

Spline merupakan salah satu dari model regresi nonparametrik. Spline merupakan polinomial truncated, yaitu bentuk kurva yang terpotong-potong sehingga spline mampu mengatasi perubahan data pada sub interval tertentu. Wahba (1990) menunjukkan bahwa spline mempunyai sifat-sifat statistik yang berguna untuk menganalisis hubungan dalam regresi. Spline dalam regresi nonparametrik terus berkembang sampai pada model *adaptive* dan multivariate respon. Contoh pendekatan regresi nonparametrik secara *adaptive* yang banyak diminati adalah *Regression Tree*, *Recursive Partitioning Regression* (RPR), dan *Multivariate Adaptive Regression Spline* (MARS) (Breiman, Olshen, & Stone, 1993).

Model nonparametrik yang dapat digunakan untuk menjelaskan hubungan kompleks dari beberapa variabel prediktor terhadap variabel respon adalah *Multivariate Adaptive Regression Spline* (MARS) (Tehupuring, 2014). *Multivariate Adaptive Regression Spline* (MARS) adalah salah satu model regresi nonparametrik yang tidak mengasumsikan bentuk hubungan fungsional antara

variabel respon dan variabel prediktor, dan mempunyai bentuk hubungan fungsional yang fleksibel. MARS merupakan pengembangan dari pendekatan *Recursive Partitioning Regression* (RPR) yang dikombinasikan dengan metode *spline* sehingga model yang dihasilkan kontinu pada *knots* yaitu garis regresi selalu menyambung, dimana tiap *knots* selalu menyambung dengan fungsi basisnya. MARS baik digunakan ketika variabel prediktor yang digunakan berjumlah banyak dan data yang digunakan bersifat tidak linear (Munoz dan Felicimo, 2009). Model MARS disusun pada pengaturan beberapa koefisien fungsi basis yang secara keseluruhan dikendalikan pada data regresi. Model MARS berguna untuk mengatasi permasalahan data berdimensi tinggi yang dikenal dengan *curse of dimensionality* dengan variabel prediktor yang bisa berinteraksi dan menghasilkan prediksi respon yang akurat serta mengatasi kelemahan regresi partisi rekursif (RPR) yaitu menghasilkan model yang kontinu pada knot, yang didasarkan pada nilai *Generalized Cross Validation* (GCV) minimum.

Penyakit diare masih menjadi masalah kesehatan masyarakat di berbagai negara, salah satunya Indonesia, karena angka kesakitan dan angka kematian dari penyakit ini tergolong tinggi. Menurut *World Health Organization* (WHO), penyakit diare adalah penyebab utama kematian kedua pada anak di bawah lima tahun. Secara global setiap tahunnya ada sekitar 2 miliar kasus diare dengan angka kematian 1,5 juta per tahun. Pada negara berkembang, anak-anak usia di bawah 3 tahun rata-rata mengalami 3 episode diare per tahun. Setiap episodenya, diare akan menyebabkan kehilangan nutrisi yang dibutuhkan anak untuk tumbuh, sehingga diare merupakan penyebab utama malnutrisi pada anak.



Indonesia merupakan salah satu negara berkembang dengan angka kejadian penyakit diare yang tinggi karena tingginya morbiditas dan mortalitas (Magdarina, 2010). Penyakit diare merupakan penyakit endemis dan juga merupakan penyakit yang berpotensi Kejadian Luar Biasa (KLB) yang disertai dengan kematian (KEMENKES RI, 2018). Berdasarkan Survey morbiditas diare pada tahun 2014 insiden diare pada balita yaitu 27% dan tahun 2016 diperkirakan jumlah penderita sebanyak 46,4% (KEMENKES RI, 2016). Target SDGs pada tahun 2030 mengakhiri kematian bayi dan balita dengan upaya mengurangi angka kematian bayi dengan 12/1000 kelahiran hidup dan angka kematian anak bawah lima tahun 25/1000 kelahiran hidup (KEMENKES RI, 2015).

Berdasarkan data pada profil kesehatan Indonesia tahun 2019 diketahui angka kesakitan diare per 1000 penduduk pada usia balita sebesar 843 kasus. Pemerintah telah menetapkan kebijakan dalam menurunkan angka kesakitan dan angka kematian karena diare, diantaranya adalah melaksanakan tatalaksana penderita diare yang sesuai standar, baik di sarana kesehatan maupun di rumah tangga, melaksanakan surveilans epidemiologi dan pengendalian Kejadian Luar Biasa, meningkatkan pengetahuan dan keterampilan petugas dalam pengelolaan program yang meliputi aspek manajerial dan teknis medis, melaksanakan evaluasi sebagai

dasar perencanaan selanjutnya.

Provinsi Jawa Barat dan Jawa Tengah merupakan dua provinsi yang angka kesakitan diare terbilang cukup tinggi. Berdasarkan data dari Kementerian Kesehatan Republik Indonesia pada tahun 2019, tercatat kasus penemuan diare pada balita di Jawa Barat sebesar 729.500 kasus. Sedangkan di Jawa Tengah, kasus penemuan diare pada balita sebesar 442.854 kasus. Angka tersebut jauh dari salah satu langkah dalam target pencapaian *Millennium Development Goals* (MDGS) *Goal* ke-4 yaitu menurunkan kematian anak menjadi 2/3 bagian dari tahun 1990 sampai pada 2015.

Penyebab diare yaitu faktor infeksi (bakteri, virus, parasit), gangguan penyerapan makanan dan minuman di usus seperti penyerapan karbohidrat, lemak dan protein, faktor makanan seperti makanan basi, beracun, alergi terhadap makanan, faktor psikologis seperti cemas, takut dan terkejut (Brand et.al, 2015). Penyebab lain dari diare adalah rotavirus, kualitas air minum, kebersihan dan sanitasi (Gul R, Hussain, Ali W, et al, 2017).

Penyebaran penyakit diare dapat terjadi secara langsung maupun tak langsung. Diare dapat ditularkan dari orang satu ke orang lain secara langsung melalui *fecal-oral* dengan media penularan utama adalah makanan dan minuman yang terkontaminasi agen penyebab diare (Suryono, 1991). Penularan penyakit diare tersebut juga bisa disebabkan karena faktor kebersihan dari lingkungan sekitar. Untuk menekan angka penyebaran diare perlu dilakukan analisis pada faktor-faktor tersebut sehingga dapat dilakukan tindakan pencegahan terhadap penyakit diare.

Oleh karena itu, berdasarkan penelitian sebelumnya maka model MARS cocok digunakan dalam data kasus kejadian diare dan faktor-faktor yang mempengaruhinya dengan pendekatan regresi nonparametrik dan data berimensi tinggi mampu mengolah data. Variabel respon dalam penelitian ini adalah jumlah kasus kejadian diare pada balita. Sedangkan variabel prediktornya adalah Tempat Pengelolaan Makanan (TPM) yang memenuhi syarat kesehatan, jumlah air minum yang memenuhi syarat kesehatan, dan pemberian ASI eksklusif pada bayi usia kurang dari 6 bulan. Berdasarkan uraian di atas, penulis mengambil judul “*Multivariate Adaptive Regression Spline (MARS) untuk Kasus Penyakit Diare Pada Balita di Provinsi Jawa Barat dan Jawa Tengah tahun 2019*”.

## **1.2 Rumusan Masalah**

Berdasarkan uraian dari latar belakang sebelumnya maka dapat ditentukan rumusan masalah dalam penelitian ini antara lain:

1. Bagaimana langkah-langkah pemodelan menggunakan metode *Multivariate Adaptive Regression Spline (MARS)*?
2. Bagaimana hasil pemodelan terbaik untuk kasus kejadian diare pada balita di Provinsi Jawa Barat-Jawa Tengah dari metode *Multivariate Adaptive Regression Spline (MARS)* menggunakan (*Generalized Cross Validation*) GCV terkecil?

## **1.3 Tujuan Penelitian**

Adapun tujuan penelitian yang ingin dicapai dalam penelitian ini antara lain:

1. Mengetahui langkah-langkah pemodelan menggunakan metode Multivariate Adaptive Regression Spline (MARS).
2. Memperoleh model kasus kejadian diare pada balita di Provinsi Jawa Barat-Jawa Tengah dengan menggunakan Multivariate Adaptive Regression Spline (MARS).

#### **1.4 Batasan Masalah**

Batasan masalah yang dilakukan pada penelitian ini adalah sebagai berikut.

1. Data yang digunakan pada penelitian ini adalah data sekunder dari Dinas Kesehatan Provinsi Jawa Barat dan Jawa Tengah tahun 2019.
2. Penelitian ini hanya menggunakan metode *Multivariate Adaptive Regression Spline* (MARS).

#### **1.5 Manfaat Penelitian**

Manfaat yang diperoleh dalam penulisan skripsi antara lain sebagai berikut:

##### **1.5.1 Bagi Mahasiswa**

1. Mahasiswa memperoleh pengetahuan tentang metode *Multivariate Adaptive Regression Spline* (MARS).
2. Mahasiswa memperoleh pengetahuan tentang prosedur penggunaan metode *Multivariate Adaptive Regression Spline* (MARS).
3. Mahasiswa memperoleh pengetahuan tentang bagaimana kasus kejadian diare dengan menggunakan metode *Multivariate Adaptive Regression Spline* (MARS).

### 1.5.2 Bagi Pembaca

1. Dapat menambah atau memperkaya khasanah kepustakaan Jurusan Matematika.
2. Menambah topik kajian tentang metode *Multivariate Adaptive Regression Spline* (MARS).

### 1.6 Tinjauan Pustaka

Penelitian sebelumnya yang berkaitan dengan *Multivariate Adaptive Regression Spline* (MARS) sebagai berikut:

1. Jurnal penelitian yang berjudul “Analisis Regresi Nonparametrik Spline Linear” oleh Fitriana Nurul Hidayah Jurusan Matematika Fakultas Sains dan Teknologi, Universitas Islam Negeri Sunan Kalijaga Yogyakarta. Penelitian ini menjelaskan tentang bagaimana memodelkan indikator-indikator yang berpengaruh pada Indeks Pembangunan Manusia pada tahun 2017 di Indonesia. Hasil studi kasus dalam penelitian ini menunjukkan model terbaik dengan menggunakan orde dua dengan satu titik knot dan model regresi *spline* linear lebih baik daripada model regresi linear, yang dapat dilihat dari koefisien determinasi dari regresi *spline* linear yang lebih besar yaitu 0,9995 dan nilai MSE yang minimum yaitu 0,0081.
2. Jurnal penelitian yang berjudul “Pemodelan Kemiskinan di Kabupaten Jombang dengan Pendekatan *Multivariate Adaptive Regression Splines* (MARS)” oleh Milliatur Rodliyah, Santi Wulan Purnami, dan Bambang Widjanarko Otok Jurusan Statistika Fakultas Matematika dan Ilmu

Pengetahuan Alam, Institut Teknologi Sepuluh November (ITS). Penelitian ini menjelaskan tentang bagaimana memodelkan kemiskinan di Kabupaten Jombang sehingga diperoleh faktor-faktor yang mempengaruhi kemiskinan di Kabupaten Jombang dengan menggunakan metode MARS. Hasil penelitiannya menunjukkan bahwa model MARS terbaik adalah kombinasi  $BF=72$ ,  $MI=2$ , dan  $MO=1$  dengan nilai CGV sebesar 26,835.

3. Jurnal penelitian yang berjudul “Pemodelan Kasus Penyakit Diare Pada Balita di Kota Surabaya Menggunakan Regresi Nonparametrik Spline Truncated” oleh Muhammad Syauqi Khudzaifi jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh November pada tahun 2017. Penelitian ini menunjukkan bagaimana memodelkan kasus penyakit diare dengan menggunakan regresi nonparametrik spline truncated. Berdasarkan hasil analisis dari penelitian tersebut adalah keempat faktor atau variabel yang digunakan dalam penelitian berpengaruh signifikan terhadap presentase kasus diare balita di Kota Surabaya dan model regresi nonparametrik spline truncated menghasilkan koefisien determinasi sebesar 91,04%.

**Tabel 1.1 Tinjauan Pustaka**

No	Peneliti	Metode	Estimasi	Objek Penelitian
1.	Fitriana Nurul Hidayah (2019)	Regresi Spline dengan Metode <i>Mean Square Error</i> (MSE) dan <i>Generalized Cross Validation</i> (GCV)	<i>Ordinary Least Square</i>	Data Indeks Pembangunan Manusia
2.	Milliatur Rodliyah, Santi Wulan Purnami, dan Bambang Widjanarko Otok (2014)	<i>Multivariate Adaptive Regression Splines</i> (MARS)	<i>Penalized Least Square</i> (PLS)	Pemodelan Kemiskinan di Kabupaten Jombang

**Tabel 1.2 Tinjauan Pustaka**

No	Peneliti	Metode	Estimasi	Objek Penelitian
3.	Muhammad Syauqi Khudzaifi (2017)	Regresi Nonparametrik <i>Spline</i> <i>Truncated</i>	<i>Ordinary</i> <i>Least</i> <i>Square</i> (OLS)	Kasus Penyakit Diare Pada Balita di Kota Surabaya Tahun 2015
4.	A'yun Nafsi Utami	<i>Multivariate</i> <i>Adaptive</i> <i>Regression</i> <i>Spline</i> (MARS) Dengan Metode <i>Generalized</i> <i>Validation Cross</i> (GCV)	<i>Maximum</i> <i>Likelihood</i>	Kasus Kesakitan Diare Pada Balita di Provinsi Jawa Barat dan Jawa Tengah Tahun 2019

Persamaan penelitian ini akan membahas mengenai *Multivariate Adaptive Regression Spline* (MARS) dengan pemilihan model terbaik menggunakan nilai GCV minimum atau terkecil. Perbedaan penelitian ini dengan penelitian sebelumnya adalah dalam menentukan model MARS berdasarkan nilai GCV terkecil serta dengan menggunakan estimasi *Maximum Likelihood* dan penelitian ini dikembangkan model pemrograman menggunakan software R dengan studi kasus jumlah kasus kejadian diare pada balita di Provinsi Jawa Tengah-Jawa Barat Tahun 2019.



### 1.7 Sistematika Penulisan

Penelitian ini terdiri dari beberapa bab untuk mempermudah pemahaman dan pembahasan terhadap permasalahan yang diangkat dan diteliti, secara umum dapat dilihat dari sistematika penulisan di bawah ini:

#### BAB I PENDAHULUAN

Pada bagian ini merupakan bab pendahuluan yang berisi tentang Latar Belakang, Rumusan Masalah, Tujuan Penelitian, Manfaat Penelitian, Batasan Masalah, Tinjauan Pustaka, dan Sistematika Penulisan.

#### BAB II Landasan Teori

Pada bagian ini akan membahas mengenai teori-teori yang akan digunakan sebagai dasar dalam pemecahan masalah. Teori-teori tersebut antara lain: matriks, teori probabilitas, estimasi *maximum likelihood*, analisis regresi, model regresi linier sederhana, estimasi parameter dengan *maximum likelihood*, asumsi model regresi linier sederhana, pengujian parameter model regresi, penyimpangan dalam analisis regresi, regresi nonparametrik, spline, diare.

#### BAB III Metode Penelitian

Bagian ini merupakan bab metode penelitian yang berisi tentang studi pustaka, jenis dan sumber data, variabel, *software* yang digunakan, tahapan penelitian, dan *flowchat* penelitian.

#### BAB IV Pembahasan

Pada bagian ini membahas tentang model MARS, estimasi parameter menggunakan *maximum likelihood*, *Generalized Cross Validation* (GCV), pengujian signifikansi model MARS, pemilihan terbaik model MARS.

#### BAB V Studi Kasus

Pada bab ini membahas tentang deskripsi data yang digunakan serta analisis data dengan *Multivariate Adaptive Regression Spline* (MARS) menggunakan estimasi *maximum likelihood*.

#### BAB VI Penutup

Pada bagian ini berisi tentang kesimpulan-kesimpulan yang diperoleh dari beberapa pembahasan sebelumnya dan saran-saran yang berkaitan dengan penelitian skripsi ini sebagai akibat dari kekurangan atau kelebihan dari hasil analisis yang dilakukan.

#### DAFTAR PUSTAKA

Bagian ini berisi keterangan dari beberapa buku dan literatur lain yang menjadi acuan dalam penyusunan tugas akhir ini.

## BAB VI

### KESIMPULAN DAN SARAN

#### 6.1 Kesimpulan

Berdasarkan hasil analisis di atas dapat disimpulkan bahwa :

1. Langkah-langkah dalam pemodelan menggunakan metode *Multivariate Adaptive Regression Spline* (MARS) antara lain:
  - a. Melakukan analisis deskriptif pada masing-masing variabel prediktor sebagai langkah awal mengetahui pola hubungan antar variabel.
  - b. Membuat *scatterplot* antara variabel respon dengan masing-masing variabel prediktor.
  - c. Melakukan uji asumsi klasik pada variabel-variabel tersebut untuk memastikan dapat dilakukan analisis regresi nonparametrik. Apabila menyalahi atau melanggar aturan uji asumsi klasik maka dilanjutkan pada langkah selanjutnya.
  - d. Menentukan model MARS terbaik dengan berdasarkan nilai *Generalized Cross Validation* (GCV) terkecil dari *software R*.
  - e. Melakukan interpretasi model MARS terbaik dan interpretasi variabel-variabel yang berpengaruh di model.
  - f. Menentukan tingkat kepentingan masing – masing variabel prediktor yang berpengaruh.

- g. Menguji signifikansi model MARS untuk mengevaluasi kecocokan model dengan uji koefisien regresi secara simultan (Uji F) maupun secara parsial (Uji T).
2. Hasil pemilihan model terbaik pada jumlah kasus diare pada balita di Provinsi Jawa Tengah-Jawa Barat tahun 2019 dengan menggunakan metode *Multivariate Adaptive Regression Spline* (MARS) dengan nilai GCV terkecil adalah

$$\hat{Y} = -0,6236969 + 1,011132 * BF_1 - 1,308728 * BF_2 + 0,2991534 * BF_3 + 0,5829459 * BF_4$$

Dengan

$$BF_1 = \max(0, zx1 - 0,03701)$$

$$BF_2 = \max(0, zx1 - 0,89081)$$

$$BF_3 = \max(0, zx2 - 0,43872)$$

$$BF_4 = \max(0, zx3 - 0,64859)$$

Variabel-variabel prediktor yang tingkat pentingnya terhadap model terbaik yang diperoleh secara signifikan mempengaruhi kasus diare adalah pemberian ASI eksklusif ( $X_3$ ) dengan tingkat kepentingan sebesar 100%, air minum sehat ( $X_2$ ) dengan tingkat kepentingan sebesar 44,1%, dan tempat pengelolaan makanan ( $X_1$ ) dengan tingkat kepentingan sebesar 30,2%.

## 6.2 Saran

Berdasarkan hasil penelitian dan kesimpulan di atas, peneliti memberikan saran sebagai berikut:

1. Variabel-variabel prediktor yang ada di penelitian ini dapat ditambahkan oleh peneliti lanjutan dengan variabel-variabel prediktor lainnya yang dapat mempengaruhi jumlah kasus diare pada balita.
2. Dalam menentukan model terbaik peneliti lanjutan bisa dengan menggunakan metode selain nilai GCV terkecil seperti MSE atau  $R^2$ .

## DAFTAR PUSTAKA

- Anton, H. (1991). *Aljabar Linear Elementer Edisi Kelima*. Jakarta: Penerbit Erlangga.
- Arifin, M. Z., & Mitakda, M. B. (2015). Metode Pendekatan Multivariate Adaptive Regression Spline (MARS) Bagging Dalam Memodelkan Presentase Gizi Buruk Balita di Jawa Timu. *Jurnal Mahasiswa Statistik Vol 3*.
- Asriani, E. D. (2016). *Estimasi Multivarite Adaptive Regression Spline (MARS) Pada Indeks Harga Saham (IHSG)*. Semarang: Universitas Negeri Semarang.
- Awat, N. J. (1995). *Metode Statistika dan Ekonometri*. Liberty, Yogyakarta.
- Baskoro, A. A. (2019). Model Regresi Nonparametrik Spline Kuadrat dalam Skripsi Fakultas Sains dan Teknologi UIN Sunan Kalijaga.
- Basri, H. (2008). Estimasi Kurva Regresi Nonparametrik dengan Pedekatan Spline. *Jurnal Kependidikan Vol 3*.
- Breiman, L., Friedman, J., Olshen, R., & Stone, J. (1993). *Classification and Regression Trees*. New York: Chapman and Hall.
- Budiantara, S. F. (2012). Analisis Survival dengan Pendekatan Multivariate Adaptive Regression Spline pada Kasus Demam Berdarah Dengue (DBD). *Jurnal Sains dan Seni, 1(1)*, 318-232.
- Damodar, G. d. (1995). *Basic Econometrics. Third Edition*. Singapura: McGraw-Hill.
- Draper, N. d. (1992). *Analisis Reregresi Terapan, Terjemahan Bambang Sumantri*. Jakarta: Gramedia.
- Eubank, R. (1988). *Spline Smoothing and Nonparametric Regression*. New York: Marcel Dekker.
- Eubank, R. (1999). *Nonparametric Regression and Spline Smoothing*. Second Edition, Marcel Dekker, New York.

- Friedman, J. (1991). Multivariate Adaptive Regression Spline (With Discussion) .  
*The Annals of Statistic*, Vol. 19, hal 1-141.
- Ghofar, R. Y., Safitri, D., & Rusdiyono, A. (2014). Klasifikasi Kelulusan Mahasiswa Fakultas Sains dan Matematika Universitas Diponegoro Menggunakan Multivariate Adaptive Regression Spline (MARS). *Jurnal Gaussian Vol 3*.
- Gujarati, D. N., & Porter, D. C. (2010). *Dasar-Dasar Ekonometrika Edisi 5*. Jakarta: Salemba Empat.
- Hardle, W. (1990). *Aplied Nonparametric Regression*. Cambridge University.
- Harini, S., & Turmudi. (2008). *Metode Statistika*. Malang: UIN Malang Press.
- Hasan, M. I. (2002). *Pokok-Pokok Materi Metodologi Penelitian dan Aplikasinya*. Jakarta: Ghalia Indonesia.
- <http://www.diskes.jabarprov.go.id/>. (2020, 12 15).
- <https://dinkesjatengprov.go.id/>. (2020, 12 15).
- Imrona, M. (2009). *Aljabar Linear Dasar*. Jakarta: Erlangga.
- J. Neter, W. W. (1997). *Analisis Regresi Linear Sederhana. Alih Bahasa Bambang Sumantri*. Bogor: FMIPA IPB.
- Johannes, & Budiono. (1994). *Pengantar Matematika untuk Ekonomi*. Jakarta: LP3ES.
- Kasse, I. (2017). Memodelkan Faktor-Faktor Yang Mempengaruhi Gizi Buruk Balita Dengan Metode Multivariate Adaptive Regression Spline (MARS). *Matematika dan Statistika serta Aplikasinya Vol 5, No 1*.
- Kishartini, S. I. (2014). Multivariate Adaptive Regression Spline (MARS) untuk Klasifikasi Status Kerja di Kabupaten Demak. *Jurnal Gaussian*, 3(4):711-718.
- Kriner, M. (2007). *Survival Analysis with Multivariate Adaptive Regression Spline*. German. Munchen University.
- Kutner, M. H., Nachtsheim, C. J., & Neter, J. (2004). *Applied Linear Regression Models*. New York: Mc Graw-Hill Companies, Inc.

- M. F. Qudratullah, S. U. (2012). *Statistika*. Yogyakarta: SUKA-Press UIN Sunan Kalijaga.
- Meliala, B. (2017). Regresi Spline Multivariat Adaptive (Multivariate Adaptive Regression Spline). *Jurnal Matematika*.
- Otok, B. W. (2008). Asimtotik Model Mutivariate Adaptive Regression Spline. *Jurnal S3 Matematika, 10(2)*, 112-119.
- Otok, B. W. (2010). Pendekatan Multivariate Adaptive Regression Spline (MARS) pada Pengelompokan Zona Musim Suatu Wilayah. *STATISTIKA: Forum Teori dan Aplikasi Statistik Vol 10*.
- Otok, W. P. (2012). Pemodelan Kemiskinan di Provinsi Jawa Timur dengan Pendekatan Multivariate Adaptive. *Jurnal Sains dan Seni ITS Vol. 1 No. 1*.
- Qudratullah, M. F. (2013). *Analisis Regresi Terapan Teori, Contoh Kasus dan Aplikasi dengan SPSS*. Yogyakarta: Andi Offset.
- Raditya, A. P. (2018). *Penerapan Metode Multivariate Adaptive Regression Spline (MARS) Dengan Estimator Nadaraya-Watson Fungsi Kernel Gaussian*. Skripsi.
- Rahmadhani, N. F. (2019). *Multivariate Adaptive Regression Spline (MARS) dalam Meenentukan Faktor-Faktor Kecelakaan Lalu Lintas di Kota Yogyakarta dalam Skripsi UIN Sunan Kalijaga Yogyakarta*.
- Suliyanto. (2014). *STATISTIK NON PARAMETRIK dalam Aplikasi Penelitian*. Yogyakarta: Penerbit Andi.
- Dinas Kesehatan Provinsi Jawa Tengah. (2020). *Profil Kesehatan Provinsi Jawa Tengah Tahun 2019*.
- Dinas Kesehatan Provinsi Jawa Barat. (2020). *Profil Kesehatan Provinsi Jawa Barat Tahun 2019*.
- Wahba, G. (1990). *Spline Models for Observasional Data*. Pennsylvania. SIAM.
- Walpole, R. E. (1993). *Pengantar Statistika Edisi ke-3*. Jakarta: PT. Gramedia Pustaka Utama.



Walpole, R. E., & Myers, R. H. (1995). *Ilmu Peluang dan Statistika unuk Insyinyur dan Ilmuwan Edisi ke-4*. Bandung: Penerbit ITB.

Wicaksono, W. (2014). *Pemodelan Multivariate Adaptive Regression Splines (MARS) pada Faktor-Faktor Resiko Angka Kesakitan Diare*. Semarang: FSM Universitas Diponegoro.

## LAMPIRAN

### Lampiran 1. Tabel Data Penelitian

No	Y	X1	X2	X3
1	10181	1700	18582	10429
2	27382	3863	838	7240
3	23614	787	3081	9273
4	21041	1497	2482	818
5	36208	1943	13358	9603
6	10882	459	106	3607
7	23745	379	177	8845
8	10341	1105	6929	6089
9	8156	1234	320	4913
10	17436	1509	766	12341
11	9586	891	240	9158
12	1639	1021	427	2782
13	3388	392	564	6673
14	10227	3028	4160	4227
15	7669	459	592	7563
16	5649	561	350	3382
17	10616	1438	324	5858
18	7232	3412	121	11800

<b>No</b>	<b>Y</b>	<b>X1</b>	<b>X2</b>	<b>X3</b>
19	3037	917	7425	4140
20	9099	1434	310	6324
21	19750	1872	7527	6122
22	11217	1583	298	3683
23	10731	921	501	5166
24	6164	1586	3747	8516
25	11490	337	126	5967
26	30183	700	62	4556
27	13331	816	214	6995
28	44421	558	121	1209
29	12935	611	68	18351
30	2316	536	76	4425
31	8038	813	671	4207
32	3229	264	149	1016
33	31431	1933	3144	18523
34	10299	606	155	2938
35	15685	222	52	742
36	138853	2747	20286	30630
37	85337	2807	7033	23685
38	64129	2494	7782	18635

<b>No</b>	<b>Y</b>	<b>X1</b>	<b>X2</b>	<b>X3</b>
39	76247	3517	647	30540
40	65157	726	0	37803
41	21838	5187	216	24362
42	28835	55	0	8389
43	27959	1598	606	13200
44	55890	1595	327	40221
45	33380	951	379	15708
46	19507	3374	276	10868
47	34746	7675	345	7400
48	33286	1289	210	12031
49	29622	1128	52	7510
50	67876	2056	624	8391
51	38433	4303	193	24384
52	26899	1501	78	16901
53	5951	108	0	5111
54	33571	2366	113	2894
55	13983	102	261	4073
56	61772	2159	159	21873
57	18567	980	71	5416
58	22126	2103	346	8827

<b>No</b>	<b>Y</b>	<b>X1</b>	<b>X2</b>	<b>X3</b>
59	20586	1360	0	13893
60	10143	978	29	1966
61	13390	211	677	5770
62	5370	455	323	2413

**Lampiran 2. Tabel Data Penelitian Setelah Standarisasi**

<b>No</b>	<b>ZY</b>	<b>ZX1</b>	<b>ZX2</b>	<b>ZX3</b>
1	-0,60307	0,12084	4,03745	0,02955
2	0,10324	1,71152	-0,26696	-0,33064
3	-0,05148	-0,55058	0,27716	-0,10102
4	-0,15713	-0,02844	0,13185	-1,056
5	0,46565	0,29955	2,77019	-0,06374
6	-0,57428	-0,79179	-0,44453	-0,74099
7	-0,0461	-0,85063	-0,4273	-0,14936
8	-0,5965	-0,31672	1,21062	-0,46065
9	-0,68622	-0,22186	-0,39261	-0,59347
10	-0,30516	-0,01962	-0,28442	0,24551
11	-0,6275	-0,4741	-0,41202	-0,114
12	-0,95382	-0,3785	-0,36666	-0,83417
13	-0,882	-0,84107	-0,33342	-0,39468

<b>No</b>	<b>ZY</b>	<b>ZX1</b>	<b>ZX2</b>	<b>ZX3</b>
14	-0,60118	1,09746	0,53891	-0,67096
15	-0,70622	-0,79179	-0,32663	-0,29416
16	-0,78916	-0,71678	-0,38534	-0,7664
17	-0,58521	-0,07183	-0,39164	-0,48674
18	-0,72416	1,37986	-0,44089	0,18441
19	-0,89641	-0,45498	1,33094	-0,68078
20	-0,6475	-0,07477	-0,39504	-0,4341
21	-0,21015	0,24733	1,35569	-0,45692
22	-0,56053	0,0348	-0,39795	-0,7324
23	-0,58048	-0,45204	-0,34871	-0,5649
24	-0,76801	0,03701	0,43872	-0,18652
25	-0,54932	-0,88151	-0,43968	-0,47443
26	0,21825	-0,61456	-0,4552	-0,6338
27	-0,47372	-0,52925	-0,41833	-0,35831
28	0,8029	-0,71899	-0,17671	-1,01184
29	-0,48998	-0,68001	-0,45374	0,92434
30	-0,92602	-0,73517	-0,4518	-0,64859
31	-0,69106	-0,53146	-0,30747	-0,67322
32	-0,88853	-0,9352	-0,4341	-1,03364
33	0,2695	0,29219	0,29244	0,94377

<b>No</b>	<b>ZY</b>	<b>ZX1</b>	<b>ZX2</b>	<b>ZX3</b>
34	-0,59822	-0,68369	-0,43264	-0,81655
35	-0,37706	-0,96608	-0,45763	-1,06459
36	4,68046	0,89081	4,45081	2,31125
37	2,48299	0,93494	1,23585	1,52681
38	1,61214	0,70476	1,41755	0,95642
39	2,10973	1,45707	-0,31329	2,30108
40	1,65436	-0,59544	-0,47024	3,12143
41	-0,12441	2,6852	-0,41784	1,60328
42	0,1629	-1,0889	-0,47024	-0,20086
43	0,12693	0,04583	-0,32323	0,34254
44	1,27384	0,04363	-0,39092	3,39454
45	0,34953	-0,42997	-0,3783	0,62581
46	-0,22012	1,35191	-0,40329	0,07914
47	0,40562	4,51489	-0,38655	-0,31257
48	0,34567	-0,18141	-0,4193	0,2105
49	0,19522	-0,29981	-0,45763	-0,30014
50	1,766	0,38265	-0,31887	-0,20064
51	0,55702	2,0351	-0,42342	1,60576
52	0,08341	-0,0255	-0,45132	0,76056
53	-0,77676	-1,04992	-0,47024	-0,57111

No	ZY	ZX1	ZX2	ZX3
54	0,35737	0,61062	-0,44283	-0,82152
55	-0,44695	-1,05433	-0,40693	-0,68835
56	1,51536	0,45839	-0,43167	1,32215
57	-0,25872	-0,40865	-0,45302	-0,53666
58	-0,11258	0,41721	-0,38631	-0,15139
59	-0,17582	-0,12919	-0,47024	0,42081
60	-0,60463	-0,41012	-0,46321	-0,92634
61	-0,4713	-0,97417	-0,30601	-0,49668
62	-0,80062	-0,79474	-0,39189	-0,87585

### Lampiran 3. Script dan Output Software R Boxplot

```
> summary(x1)
      Y           X1           X2           X3
Min.   : 1639   Min.   : 55.0   Min.   : 0.0   Min.   : 742
1st Qu.: 10152  1st Qu.: 607.2   1st Qu.: 131.8  1st Qu.: 4276
Median : 18002  Median :1181.0   Median : 323.5  Median : 7320
Mean   : 24868  Mean   :1535.7   Mean   : 1938.5  Mean   :10167
3rd Qu.: 31119  3rd Qu.:1940.5   3rd Qu.: 743.8  3rd Qu.:12264
Max.   :138853  Max.   :7675.0   Max.   :20286.0  Max.   :40221

> Boxplot( ~ X1, data=x1, id=list(method="y"))
[1] "41" "47" "51"

> Boxplot( ~ Y, data=x1, id=list(method="y"))
```



### Lampiran 4. Output SPSS Hasil Uji Simultan (Serentak)

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	35,176	3	11,725	26,335	,000 <sup>b</sup>
	Residual	25,824	58	,445		
	Total	61,000	61			

a. Dependent Variable: Zscore(y)

b. Predictors: (Constant), Zscore(x3), Zscore(x2), Zscore(x1)

### Lampiran 5. Output SPSS Hasil Uji Parsial

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3,919E-17	,085		,000	1,000
	Zscore(x1)	,095	,092	,095	1,035	,305
	Zscore(x2)	,275	,087	,275	3,150	,003
	Zscore(x3)	,615	,093	,615	6,642	,000

a. Dependent Variable: Zscore(y)

### Lampiran 6. Script dan Output *Software* R Uji Normalitas

```
> library(readxl)
> satu<-read_xlsx("E:/kuliah/MARS/satu.xlsx")
> view(satu)
> model=lm(zy~zx1+zx2+zx3,data=satu)
> library(normtest)
> library(nortest)
> lillie.test(model$residuals)

      Lilliefors (Kolmogorov-Smirnov) normality
      test

data:  model$residuals
D = 0.11366, p-value = 0.04507
```

### Lampiran 7. Script dan Output *Software R Uji Heterokedastisitas*

```

> library(lmtest)
Loading required package: zoo
Attaching package: 'zoo'

The following objects are masked from 'package:base':

  as.Date, as.Date.numeric

warning messages:
1: package 'lmtest' was built under R version 3.6.3
2: package 'zoo' was built under R version 3.6.3
> bptest(model)

      studentized Breusch-Pagan test

data:  model
BP = 20.838, df = 3, p-value = 0.0001137
>

```

### Lampiran 8. Script dan Output *Software R Uji Autokorelasi*

```

> dwtest(model)

      Durbin-watson test

data:  model
DW = 1.5303, p-value = 0.02422
alternative hypothesis: true autocorrelation is greater
than 0
> |

```

### Lampiran 9. Script dan Output *Software R Uji Multikolinearitas*

```

> library(car)
> vif(model)
      zx1      zx2      zx3
1.155457 1.045700 1.173026
>

```

## Lampiran 10. Script dan Output *Software R Model MARS*

```

> library(TeachingDemos)
> library(earth)
> mars<-earth(zy~zx1+zx2+zx3,data=satu,pmethod="backward",nprune = 30,nfold = 20)
> summary(mars,digit=3)
Call: earth(formula=zy~zx1+zx2+zx3, data=satu,
             pmethod="backward", nprune=30, nfold=20)

             coefficients
(Intercept)      -0.624
h(zx1-0.03701)    1.011
h(zx1-0.89081)   -1.309
h(zx2-0.43872)   0.299
h(zx3- -0.64859) 0.583

Selected 5 of 9 terms, and 3 of 3 predictors (nprune=30)
Termination condition: RSq changed by less than 0.001 at 9 terms
Importance: zx3, zx2, zx1
Number of terms at each degree of interaction: 1 4 (additive model)
GCV 0.475  RSS 21.5  GRSq 0.533  RSq 0.648  CVRSq -22.3

Note: the cross-validation sd's below are standard deviations across folds

Cross validation:  nterms 4.70 sd 0.80  nvars 2.80 sd 0.41

             CVRSq  sd  MaxErr  sd
             -22.3 92.1    4.24 1.52
> earth.mod<-earth(zy~zx1+zx2+zx3,data=satu)
> cat(format(mars,style = "max"))
-0.6236969
+ 1.011132 * max(0,      zx1 - 0.03701)
- 1.308728 * max(0,      zx1 - 0.89081)
+ 0.2991534 * max(0,      zx2 - 0.43872)
+ 0.5829459 * max(0,      zx3 - -0.64859)
> |

```

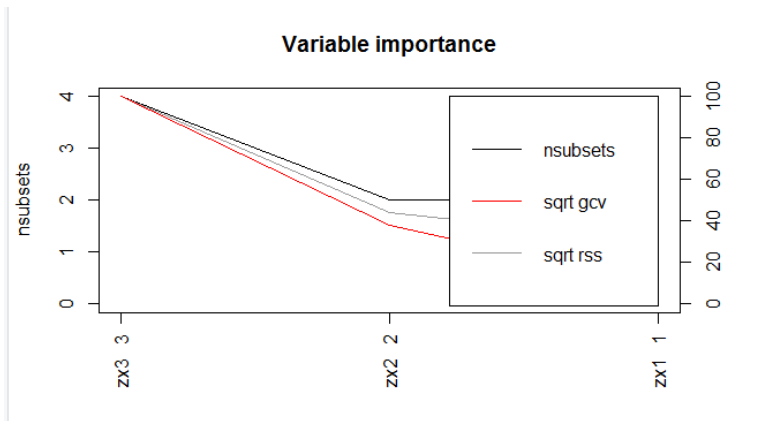
## Lampiran 11. Script dan Output *Software R Tingkat Kepentingan*

```

> evimp(mars,trim=TRUE,sqrt. = TRUE)
      nsubsets  gcv  rss
zx3         4 100.0 100.0
zx2         2  38.1  44.1
zx1         2   9.0  30.2

> ev<-evimp(mars,trim=FALSE)
> print(ev)
      nsubsets  gcv  rss
zx3         4 100.0 100.0
zx2         2  38.1  44.1
zx1         2   9.0  30.2
> plot(ev)
> |

```



### Lampiran 12. Script dan Output Software R Uji Simultan

```
> anova(model)
Analysis of Variance Table

Response: zy
      Df Sum Sq Mean Sq
zx1    1  7.7136   7.7136
zx2    1  7.8198   7.8198
zx3    1 19.6427  19.6427
Residuals 58 25.8238  0.4452
      F value    Pr(>F)
zx1    17.325 0.0001057 ***
zx2    17.563 9.600e-05 ***
zx3    44.117 1.167e-08 ***
Residuals
----
Signif. codes:
  0 '***' 0.001 '**' 0.01 '*'
  0.05 '.' 0.1 ' ' 1
> |
```

### Lampiran 13. Script dan Output Software R Uji Parsial

```
> summary(model)

Call:
lm(formula = zy ~ zx1 + zx2 + zx3, data = satu)

Residuals:
    Min       1Q   Median       3Q      Max
-1.74393 -0.33137 -0.07848  0.24178  1.95031

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.314e-07  8.474e-02  0.000  1.00000
zx1          9.506e-02  9.184e-02  1.035  0.30491
zx2          2.752e-01  8.736e-02  3.150  0.00258 **
zx3          6.146e-01  9.253e-02  6.642  1.17e-08 ***
---
Signif. codes:
  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6673 on 58 degrees of freedom
Multiple R-squared:  0.5767,    Adjusted R-squared:  0.554
F-statistic: 26.34 on 3 and 58 DF,  p-value: 7.062e-11
>
```

## DAFTAR RIWAYAT HIDUP



### I. Data Diri

Nama : A'yun Nafsi Utami  
Tempat, tanggal lahir : Yogyakarta, 14 September 1996  
Jenis kelamin : Perempuan  
Alamat : Gonjen RT 04 Tamantirto Kasihan Bantul  
No. Hp : 0895363340127  
E-mail : ayunnaf4@gmail.com

### II. Riwayat Pendidikan

2001 – 2003 TK Pertiwi 56  
2003 – 2009 SD Negeri 2 Padokan  
2009 – 2012 SMP Negeri 2 Kasihan  
2012 – 2015 SMA Negeri 1 Sewon  
2016 – 2021 UIN Sunan Kalijaga Yogyakarta