

**PEMODELAN *MULTIVARIATE ADAPTIVE REGRESSION*
SPLINES (MARS) PADA FAKTOR-FAKTOR RESIKO
ANGKA KESAKITAN DIARE**

**(Studi Kasus : Angka kesakitan Diare di Jawa Tengah, Jawa Timur dan
Daerah Istimewa Yogyakarta Tahun 2011)**



SKRIPSI

Diajukan Sebagai Salah Satu Syarat Memperoleh Gelar

Sarjana Sains pada Jurusan Statistika FSM UNDIP

Disusun oleh

WASIS WICAKSONO

J2E009049

JURUSAN STATISTIKA

FAKULTAS SAINS DAN MATEMATIKA

UNIVERSITAS DIPONEGORO

SEMARANG

2014

**PEMODELAN *MULTIVARIATE ADAPTIVE REGRESSION*
SPLINES (MARS) PADA FAKTOR-FAKTOR RESIKO
ANGKA KESAKITAN DIARE**

**(Studi Kasus : Angka kesakitan Diare di Jawa Tengah, Jawa Timur dan
Daerah Istimewa Yogyakarta Tahun 2011)**

Diajukan Sebagai Salah Satu Syarat Memperoleh Gelar

Sarjana Sains pada Jurusan Statistika FSM UNDIP

**Disusun oleh
WASIS WICAKSONO
J2E009049**

**JURUSAN STATISTIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO
SEMARANG**

2014

HALAMAN PENGESAHAN I

Judul : Pemodelan Multivariate Adaptive Regression Splines (MARS) pada Faktor-Faktor Resiko Angka Kesakitan Diare (Studi Kasus : Angka Kesakitan Diare di Jawa Tengah, Jawa Timur dan Daerah Istimewa Yogyakarta Tahun 2011)

Nama : Wasis Wicaksono

NIM : J2E009049

Jurusan : Statistika

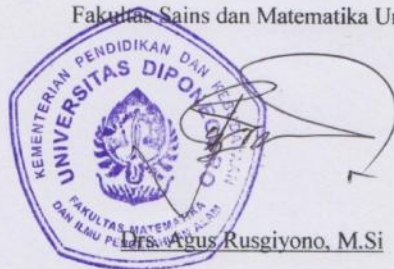
Telah diujikan pada sidang Tugas Akhir tanggal 7 Februari 2014 dan dinyatakan lulus pada tanggal 20 Februari 2014

Semarang, Februari 2014

Mengetahui,

a/n Ketua Jurusan Statistika

Fakultas Sains dan Matematika Undip



Drs. Agus Rusgiono, M.Si

NIP. 1964 08 13 1990 01 1 001

Panitia Penguji Tugas Akhir

Ketua,

Dra. Hj. Dwi Ispriyanti, M.Si

NIP. 1957 09 14 1986 03 2 001

HALAMAN PENGESAHAN II

Judul : Pemodelan Multivariate Adaptive Regression Splines (MARS) pada Faktor-Faktor Resiko Angka Kesakitan Diare (Studi Kasus : Angka Kesakitan Diare di Jawa Tengah, Jawa Timur dan Daerah Istimewa Yogyakarta Tahun 2011)

Nama : Wasis Wicaksono

NIM : J2E009049

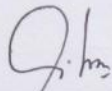
Jurusan : Statistika

Telah diujikan pada sidang Tugas Akhir pada tanggal 07 Februari 2014

Semarang, Februari 2013

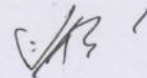
Pembimbing I

Pembimbing II



Yuciana Wilandari, S.Si, M.Si

NIP. 197005191998022001



Dra. Suparti, M.Si

NIP. 196509131990032001

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah memberikan rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan penulisan tugas akhir dengan judul **Pemodelan *Multivariate Adaptive Regression Splines* (MARS) pada Faktor-Faktor Resiko Angka Kesakitan Diare (Studi Kasus: Angka kesakitan Diare di Jawa Tengah, Jawa Timur dan Daerah Istimewa Yogyakarta Tahun 2011).**

Penulis menyadari tugas akhir ini tidak akan dapat diselesaikan tanpa bantuan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada

1. Ibu Dra. Dwi Ispriyanti, M. Si. selaku Ketua Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro.
2. Ibu Yuciana Wilandari, S.Si, M.Si selaku dosen pembimbing I.
3. Ibu Dra. Suparti, M.Si selaku dosen pembimbing II.
4. Bapak/Ibu dosen jurusan statistika yang telah memberikan arahan dan masukan demi perbaikan penulisan tugas akhir ini.
5. Semua pihak yang telah membantu dalam penulisan laporan ini.

Semarang, Februari 2014

Penulis

ABSTRAK

Angka kesakitan diare dapat diartikan sebagai jumlah penderita diare di suatu wilayah pada periode satu tahun. Tinggi rendahnya angka kesakitan diare dapat disebabkan oleh faktor lingkungan, faktor pendidikan, faktor sosial ekonomi, serta faktor gizi dan makanan. Untuk menekan angka penyebaran diare perlu dilakukan analisis pada faktor-faktor tersebut sehingga dapat dilakukan tindakan pencegahan terhadap penyakit diare. Dalam menjelaskan pola hubungan antara variabel respon dengan variabel prediktor dapat digunakan pendekatan kurva regresi. Pendekatan kurva regresi yang sering digunakan adalah pendekatan regresi parametrik, dimana bentuk kurva regresi diketahui (seperti linier, kuadratik, kubik). Jika bentuk kurva tidak diketahui maka kurva regresi dapat diduga menggunakan pendekatan model regresi nonparametrik. *Multivariate Adaptive Regression Spline* (MARS) merupakan salah satu metode regresi nonparametrik yang dapat digunakan pada data berdimensi tinggi. Untuk mendapatkan model MARS terbaik dilakukan dengan cara mengkombinasikan Minimal Observasi (MO), Maksimum Fungsi Basis (BF), dan Minimum Interaksi (MI) secara *trial and error*. Model MARS yang digunakan untuk memprediksi angka kesakitan diare di Jawa Tengah, Jawa Timur dan Daerah Istimewa Yogyakarta adalah model MARS(MO=2; BF=28; MI=3) dengan bentuk persamaan : $Y = -0.526742 + 0.264444 * BF2 + 12.2382 * BF5 - 7.76719 * BF15 + 4.96445 * BF17$

Kata Kunci : Angka Kesakitan Diare, Regresi Nonparametrik, MARS

ABSTRACT

Diarrhea morbidity can be interpreted as the number of patients with diarrhea in a region in the period of one year . Fluctuation in morbidity of diarrhea can be caused by environmental factors , educational factors , socio-economic factors , as well as nutritional and dietary factors . To reduce the number of spread of diarrhea needs to be done on the analysis of these factors are amenable to preventive measures against diarrheal diseases . To explain the relationship between the response variable and the predictor variables can be used by regression curve approach . Regression curve approach that is often used is the parametric regression approach , which assumed the form of the regression curve is known (such as linear, quadratic, cubic) . If the parametric model assumptions are not met then the regression curve can be predicted using nonparametric regression model approach . Multivariate Adaptive Regression Spline (MARS) is a nonparametric regression method that can be used in data of high dimension . To get the best MARS models is done by combining Minimal Observation (MO) , Maximum Basis Function (BF) , and the Minimum Interaction (MI) by trial and error . MARS models were used to predict morbidity of diarrhea in Central Java , East Java and Yogyakarta is a model MARS (MO = 2 ; BF = 28 ; MI = 3) in the form of the equation : $Y = -0.526742 + 0.264444 * BF2 + 12.2382 * BF5 - 7.76719 * BF15 + 4.96445 * BF17$

Keywords: Diarrhea Morbidity, Nonparametric regression, MARS

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
KATA PENGANTAR	iii
ABSTRAK	iv
ABSTRACT	v
DAFTAR ISI	vi
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Tujuan Penulisan	3
BAB II TINJAUAN PUSTAKA	
2.1 Penyakit Diare	4
2.1.1 Faktor Risiko Penyebab Diare	5
2.1.2 Pengendalian Diare di Indonesia	6
2.2 Analisis Regresi	7
2.3 Regresi Spline	12
2.4 Recursive Partitioning Regression (RPR)	13
2.5 Multivariate Adaptive Regression Splines (MARS)	15
2.5.1 Estimasi Parameter Model MARS	18
2.5.2 Dekomposisi Anova	19
2.5.3 Pemilihan Model MARS Terbaik	21
2.5.4 Pengujian Signifikansi Model MARS	22

BAB III METODOLOGI PENELITIAN

3.1 Jenis dan Sumber Data	24
3.2 Variabel Penelitian	24
3.3 Metode Analisis	25
3.3.1 Analisis Deskriptif	25
3.3.2 Analisis MARS	25
3.4 Diagram Alir Analisis	27

BAB IV ANALISIS DAN PEMBAHASAN

4.1 Deskripsi Data	28
4.2 Pemodelan Angka Kesakitan Diare Menggunakan MARS	34
4.2.1 Pemodelan MARS dengan Minimal Observasi (MO) = 0 ..	35
4.2.2 Pemodelan MARS dengan Minimal Observasi (MO) = 1 ..	36
4.2.3 Pemodelan MARS dengan Minimal Observasi (MO) = 2 ..	37
4.2.4 Pemodelan MARS dengan Minimal Observasi (MO) = 3 ..	38
4.2.5 Pemodelan MARS dengan Minimal Observasi (MO) = 4 ..	39
4.2.6 Pemodelan MARS dengan Minimal Observasi (MO) = 5 ..	40
4.3 Identifikasi Model MARS	41
4.4 Estimasi Parameter	41
4.5 Model MARS Terbaik	43
4.6 Uji Asumsi Regresi Parametrik	44
4.6.1 Uji Normalitas	44
4.6.2 Uji Homoskedastisitas	45
4.6.3 Uji Autokorelasi	47

4.7	Pengujian Signifikansi Model MARS	48
4.7.1	Pengujian Koefisien Regresi Simultan	48
4.7.2	Pengujian Koefisien Regresi Parsial	49
4.8	Interpretasi Model MARS Terbaik.....	50
4.5	Variabel-Variabel yang Berpengaruh dalam Model MARS	51
BAB V PENUTUP		
5.1	Kesimpulan.....	53
5.2	Saran.....	54
DAFTAR PUSTAKA		55

DAFTAR LAMPIRAN

Lampiran 1 Data Penelitian	57
Lampiran 2 Data Penelitian (Normal Baku).....	59
Lampiran 3 Pengolahan Data Menggunakan SPM 7.0	61
Lampiran 4 Pengolahan Data Menggunakan SPSS 16	72
Lampiran 5 Tabel Durbin-Watson	74
Lampiran 6 Tabel F	75

BAB I

PENDAHULUAN

1.1. Latar Belakang

Penyakit diare masih menjadi masalah kesehatan masyarakat di Indonesia karena angka kesakitan dan angka kematian yang ditimbulkan tergolong tinggi. Berdasarkan hasil Riset Kesehatan Dasar (Riskesdas) 2007, penyakit diare menempati urutan ke-3 penyakit menular penyebab kematian pada semua umur di Indonesia. Sedangkan angka kesakitan yang disebabkan oleh diare pada tahun 2010 mencapai 411 per 1000 penduduk. Tingginya angka kesakitan membuat penyakit diare sering menimbulkan Kejadian Luar Biasa (KLB). Jumlah kasus KLB Diare pada tahun 2010 sebanyak 2.580 dengan kematian sebesar 77 kasus sedangkan pada tahun 2009 KLB Diare sebanyak 3.037 kasus dengan kematian sebesar 21 kasus. (Dinas Kesehatan, 2010).

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 atau minuman yang terkontaminasi agen penyebab diare (Suharyono, 1991). Sedangkan penularan penyakit diare secara tidak langsung tak hanya disebabkan oleh kebersihan dan kesehatan lingkungan, tetapi juga dipengaruhi oleh faktor eksternal lainnya, seperti : faktor sosial, ekonomi dan juga pendidikan. Untuk menekan angka penyebaran diare perlu dilakukan analisis pada faktor-faktor tersebut sehingga dapat dilakukan tindakan pencegahan terhadap penyakit diare.

Analisis regresi merupakan salah satu metode statistika yang dapat menggambarkan ketergantungan atau mencari hubungan fungsional antara satu variabel respon (variabel dependen) dengan satu atau lebih variabel prediktor (variabel independen). Dalam hal ini angka kesakitan diare adalah variabel respon dan variabel prediktornya adalah faktor-faktor yang diduga sebagai penyebab tingginya angka kesakitan diare, yaitu : faktor lingkungan, faktor pendidikan, faktor sosial-ekonomi, serta faktor gizi dan makanan.

Dalam menjelaskan pola hubungan antara variabel respon dengan variabel prediktor dapat digunakan pendekatan kurva regresi. Pendekatan kurva regresi yang sering digunakan adalah pendekatan regresi parametrik, dimana diasumsikan bentuk kurva regresi diketahui (seperti linier, kuadratik, kubik) berdasarkan teori yang dapat memberikan informasi hubungan (Draper dan Smith, 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 bentuk kurva tidak diketahui maka kurva regresi dapat diduga menggunakan pendekatan model regresi nonparametrik. Regresi nonparametrik 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, 1988).

Salah satu metode regresi nonparametrik adalah *Multivariate Adaptive Regression Spline* (MARS) yang pertama kali dipopulerkan oleh Friedman (1991). Model MARS berguna untuk mengatasi permasalahan data yang berdimensi tinggi, yaitu data yang memiliki jumlah variabel prediktor sebesar $3 \leq n \leq 20$. MARS merupakan pengembangan dari pendekatan *Recursive Partition Regression* (RPR)

yang dikombinasikan dengan metode *spline* sehingga model yang dihasilkan kontinyu pada knot.

Berdasarkan uraian di atas, pokok permasalahan yang dibahas oleh penulis adalah menganalisis faktor-faktor yang mempengaruhi banyaknya angka kesakitan diare dan menggunakan metode MARS untuk mendapatkan model angka kesakitan diare yang dibatasi pada faktor-faktor yang mempengaruhi banyaknya angka kesakitan diare pada 78 kabupaten / kota di Jawa Tengah, Jawa Timur dan Daerah Istimewa Yogyakarta tahun 2011.

1.2. Tujuan Penulisan

Tujuan yang ingin dicapai pada penelitian ini adalah

1. Memperoleh model hubungan dan faktor risiko (variabel-variabel) yang diduga berpengaruh secara signifikan terhadap angka kesakitan diare di Jawa Tengah, Jawa Timur dan Daerah Istimewa Yogyakarta tahun 2011 dengan menggunakan metode MARS.
2. Mengetahui besarnya pengaruh variabel-variabel prediktor terhadap model yang diperoleh.