

MULTIVARIATE ADAPTIVE REGRESSION SPLINES (MARS)
UNTUK KLASIFIKASI STATUS KERJA
DI KABUPATEN DEMAK



SKRIPSI

Disusun oleh
KISHARTINI
24010210141008

JURUSAN STATISTIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO
SEMARANG
2014

MULTIVARIATE ADAPTIVE REGRESSION SPLINES (MARS)
UNTUK KLASIFIKASI STATUS KERJA
DI KABUPATEN DEMAK

Oleh :

KISHARTINI

24010210141008

Diajukan Sebagai Salah Satu Syarat Untuk memperoleh Gelar

Sarjana Sains pada Jurusan Statistika

JURUSAN STATISTIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO
SEMARANG

2014

HALAMAN PENGESAHAN I

Judul : *Multivariate Adaptive Regression Splines (MARS)* untuk Klasifikasi
Status Kerja di Kabupaten Demak

Nama : Kishartini

NIM : 24010210141008

Jurusan : Statistika

Telah diujikan pada sidang Tugas Akhir tanggal 16 September 2014 dan
dinyatakan lulus pada tanggal 29 September 2014

Semarang, September 2014

HALAMAN PENGESAHAN II

Judul : *Multivariate Adaptive Regression Splines* (MARS) untuk Klasifikasi
Status Kerja di Kabupaten Demak

Nama : Kishartini

NIM : 24010210141008

Jurusan : Statistika

Telah diujikan pada sidang Tugas Akhir pada tanggal 16 September 2014

Semarang, September 2014

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT, karena berkat rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul “*Multivariate Adaptive Regression Splines (MARS) untuk Klasifikasi Status Kerja di Kabupaten Demak*”.

Penulis menyadari bahwa dalam penyusunan Tugas Akhir ini tidak lepas dari bimbingan dan dukungan yang diberikan beberapa 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 dan dosen pembimbing II yang telah memberikan bimbingan serta arahan.
2. Ibu Diah Safitri, M.Si. selaku dosen pembimbing I yang telah memberikan bimbingan, arahan, dan motivasi.
3. Seluruh Dosen Jurusan Statistika FSM Universitas Diponegoro yang telah memberikan ilmu yang sangat berguna.
4. Semua pihak yang telah membantu, yang tidak dapat penulis sebutkan satu per satu. Terimakasih atas dukungan dan kerjasamanya.

Penulis menyadari masih terdapat kekurangan dalam penulisan Tugas Akhir ini. Oleh karena itu, penulis dengan sikap terbuka sangat mengharapkan kritik dan saran yang bersifat membangun dari para pembaca demi perbaikan kedepannya. Akhir kata, semoga Tugas Akhir ini dapat bermanfaat bagi semua pihak.

Semarang, September 2014

Penulis

ABSTRAK

Pengangguran merupakan salah satu permasalahan yang berkaitan dengan kegiatan perekonomian, hubungan kemasyarakatan dan juga permasalahan kemanusiaan. Pengangguran juga terjadi di Kabupaten Demak dan faktor-faktor yang diduga sebagai penyebab pengangguran di Kabupaten Demak yaitu jenis kelamin, daerah tempat tinggal, umur, status dalam rumah tangga, status perkawinan dan pendidikan. Badan Pusat Statistik Kabupaten Demak mencatat jumlah penduduk yang mencari pekerjaan (pengangguran) sebanyak 226.228 orang atau 29,55% dari jumlah penduduk usia kerja. MARS (*Multivariate Adaptive Regression Splines*) merupakan salah satu metode yang digunakan untuk pengklasifikasian. MARS digunakan untuk data yang berdimensi tinggi, yaitu data yang memiliki jumlah variabel prediktor sebesar $3 \leq n \leq 20$. Data yang digunakan pada penelitian ini adalah data sekunder dari hasil survei angkatan kerja nasional (SAKERNAS) tahun 2012. Untuk mendapatkan model MARS terbaik dilakukan dengan cara mengkombinasikan Maksimum Fungsi Basis (BF), Minimal Observasi (MO), dan Maksimum Interaksi (MI) secara *trial and error*. Model MARS yang digunakan untuk mengklasifikasikan status kerja di Kabupaten Demak adalah model MARS(BF=24, MI=3, MO=1).

Kata Kunci : Pengangguran, Klasifikasi, MARS

ABSTRACT

Unemployment is one of the issues relating to economic activities, public relations and also the problems of humanity. Unemployment also occur in Demak and factors suspected as the cause of unemployment in Demak: gender, area of residence, age, status in the household, marriage status and education. Demak BPS records the number of people looking for work (unemployed) as many as 226 228 people, or 29.55% of the working age population. MARS (Multivariate Adaptive Regression Splines) is one of the methods used for classification. MARS is used for high-dimensional data, which is data that has a number of predictor variables for $3 \leq n \leq 20$ data used in this study is a secondary data from national labor force survey (SAKERNAS) in 2012 To get the best MARS models performed with by combining Maximum Base Function (BF), Minimal Observation (MO), and Maximum Interaction (MI) by trial and error. MARS model is used to classify employment status in Demak are MARS models (BF = 24, MI = 3, MO = 1).

Keywords: Unemployment, Classification, MARS

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN I	ii
HALAMAN PENGESAHAN II	iii
KATA PENGANTAR	iv
ABSTRAK	v
ABSTRACT	vi
DAFTAR ISI	vii
DAFTAR LAMPIRAN	ix
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Perumusan Masalah	4
1.3 Pembatasan Masalah	4
1.4 Tujuan	4
BAB II TINJAUAN PUSTAKA	
2.1 Angkatan Kerja	5
2.2 Faktor-faktor Status Kerja	6
2.3 Skala Pengukuran	8
2.4 Regresi Non Parametrik	10
2.5 Regresi Spline	11
2.6 Multivariate Adaptive Regression Spline	12
2.6.1 Estimasi Parameter	13

2.6.2	Pemilihan Model MARS terbaik	14
2.6.3	Klasifikasi pada MARS	16
BAB III METODOLOGI PENELITIAN		
3.1	Jenis dan Sumber Data	20
3.2	Variabel Penelitian	20
3.3	Metode Analisis	21
3.4	Diagram Analisis	23
BAB IV ANALISIS DAN PEMBAHASAN		
4.1	Deskripsi Data	24
4.1.1	Status Kerja Menurut Jenis Kelamin	24
4.1.2	Status Kerja Menurut Umur.....	25
4.1.3	Status Kerja Menurut Status dalam Rumah Tangga.....	25
4.1.4	Status Kerja Menurut Status Perkawinan	26
4.1.5	Status Kerja Menurut Daerah Tempat Tinggal.....	26
4.1.6	Status Kerja Menurut Pendidikan.....	27
4.2	Pemodelan <i>Multivariate Adaptive Regression Splines</i> (MARS)...	27
4.3	Variabel yang Berpengaruh	35
4.4	Pengklasifikasian MARS	36
4.5	Estimasi Peluang	37
BAB V KESIMPULAN		
DAFTAR PUSTAKA		41
LAMPIRAN		43

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data Angkatan Kerja	43
Lampiran 2. Model MARS	45
Lampiran 3. Tabel Chi-Square	48

BAB I

PENDAHULUAN

1.1 Latar Belakang

Data dari BPS Kabupaten Demak (2012) tercatat penduduk Kabupaten Demak pada tahun 2012 sebesar 1.092.622 orang terdiri atas 542.879 laki-laki (49,69%) dan 549.743 perempuan (50,31%). Jumlah ini naik sebanyak 12.761 orang atau sekitar 1,18 persen dibanding tahun 2011. Seiring dengan peningkatan jumlah penduduk dari tahun ke tahun, maka penduduk usia kerja juga mengalami peningkatan. Hasil Survei Angkatan Kerja Nasional (Sakernas) pada bulan Agustus 2012 tercatat jumlah penduduk angkatan kerja Kabupaten Demak adalah 539.243 orang atau 70,45% dari keseluruhan jumlah penduduk usia kerja dan jumlah penduduk yang mencari pekerjaan (pengangguran) tercatat sebanyak 226.228 orang atau 29,55% dari jumlah penduduk usia kerja.

Angkatan kerja adalah penduduk usia kerja yang selama seminggu yang lalu mempunyai pekerjaan, baik yang bekerja maupun sementara tidak bekerja karena suatu sebab, seperti menunggu panen, pegawai yang sedang cuti dan sejenisnya. Di samping itu mereka tidak mempunyai pekerjaan tetapi sedang mencari atau mengharap pekerjaan juga masuk dalam angkatan kerja. Dalam perencanaan pembangunan, data ketenagakerjaan memegang peranan penting. Tanpa tenaga kerja program pembangunan tidaklah mungkin dapat dilaksanakan. Ketersediaan data ketenagakerjaan yang semakin lengkap dan tepat akan memudahkan pemerintah dalam membuat rencana pembangunan mengingat jumlah dan komposisi tenaga kerja selalu mengalami perubahan seiring dengan

berlangsungnya proses demografi. Penduduk angkatan kerja dibagi menjadi dua kategori yaitu bukan pengangguran dan pengangguran (Badan Pusat Statistik Kabupaten Demak, 2012).

Pengangguran adalah orang yang tidak mempunyai pekerjaan, lengkapnya orang yang tidak bekerja dan (masih atau sedang) mencari pekerjaan. Pengangguran merupakan salah satu permasalahan yang berkaitan dengan kegiatan perekonomian, hubungan kemasyarakatan dan juga permasalahan kemanusiaan (Dumairy, 1966). Faktor-faktor yang diduga mempengaruhi angkatan kerja di daerah Demak, yaitu: jenis kelamin, daerah tempat tinggal, umur, status dalam rumah tangga, status perkawinan dan pendidikan (Badan Pusat Statistik Kabupaten Demak, 2012).

Untuk menganalisis hubungan status kerja dapat digunakan analisis regresi. Analisis regresi berkenaan dengan studi ketergantungan satu variabel (variabel respon) pada satu atau lebih variabel lain (variabel prediktor), dengan maksud menaksir atau meramalkan nilai rata-rata hitung (mean) atau rata-rata (populasi) variabel respon (Gujarati, 1978). Dalam menjelaskan pola hubungan antara variabel respon dengan variabel prediktor dapat digunakan pendekatan kurva regresi. Pendekatan dalam menentukan kurva regresi dapat menggunakan dua cara yaitu pendekatan parametrik dan pendekatan nonparametrik. Apabila kurva regresi membentuk suatu pola hubungan tertentu seperti linier, kuadratik dan kubik serta memenuhi asumsi-asumsi klasik seperti normalitas, kesamaan varian, independensi ataupun nonautokorelasi maka dapat menggunakan pendekatan parametrik. Namun apabila asumsi dari pendekatan parametrik tidak terpenuhi maka pendugaan dapat dilakukan dengan pendekatan 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).

Pada penelitian sebelumnya, Pratiwi dan Zain (2014) telah melakukan pengklasifikasian pengangguran terbuka di Sulawesi Utara menggunakan metode *Classification and Regression Tree* (CART). Dalam penelitian tersebut faktor-faktor yang mempengaruhi pengangguran terbuka di provinsi Sulawesi Utara dengan menggunakan CART yaitu jenis kelamin, pendidikan terakhir, usia, status dalam rumah tangga, dan status perkawinan.

Pada penelitian kali ini akan digunakan metode *Multivariate Adaptive Regression Splines* (MARS) untuk mengklasifikasikan pengangguran di kabupaten Demak. *Multivariate Adaptive Regression Splines* (MARS) merupakan salah satu metode regresi nonparametrik. Karena MARS merupakan metode nonparametrik sehingga metode MARS tidak bergantung pada asumsi tertentu.

Model MARS menggunakan fungsi spline untuk menduga model. Dimana model regresi spline memberikan bentuk persamaan yang mempresentasikan bentuk parametrik polinomial *piecewise* (terbagi beberapa region). Selain itu MARS memiliki kelebihan untuk mengatasi permasalahan data yang berdimensi tinggi, yaitu data yang memiliki jumlah variabel prediktor sebesar $3 \leq n \leq 20$ (Friedman, 1991).

Penelitian tentang MARS pernah dilakukan oleh Mandaku dan Mandaku (2010) yang memodelkan kelulusan siswa masuk kelas akselerasi menggunakan regresi logistik dan *multivariate adaptive regression splines*. Budiantara dan Nisa'

(2012) juga melakukan penelitian tentang analisis survival dengan pendekatan *multivariate adaptive regression splines* pada kasus demam berdarah dengue (DBD)”.

1.2 Perumusan Masalah

Berdasarkan uraian latar belakang tersebut, masalah yang akan dibahas dalam penulisan ini adalah klasifikasi status kerja berdasarkan faktor-faktor yang mempengaruhi status kerja di Kabupaten Demak menggunakan metode *Multivariate Adaptive Regression Splines* (MARS).

1.3 Pembatasan Masalah

1. Metode yang digunakan dalam penulisan ini adalah *Multivariate Adaptive Regression Splines* (MARS) respon biner
2. Data yang digunakan adalah data survey angkatan kerja tahun 2012 oleh Badan Pusat Statistik Kabupaten Demak dengan klasifikasi pengangguran dan bukan pengangguran.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penulisan ini adalah:

1. Memperoleh model status kerja di Kabupaten Demak menggunakan metode MARS respon biner.
2. Mengklasifikasi status kerja di Kabupaten Demak.