

**Ketepatan Klasifikasi Status Pemberian Air Susu Ibu (ASI)
Menggunakan *Multivariate Adaptive Regression Splines* (MARS) dan
Algoritma C4.5 di Kabupaten Sragen**



SKRIPSI

**Disusun Oleh :
YUSUF ARIFKA RAHMAN
24010211120011**

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

2015

**Ketepatan Klasifikasi Status Pemberian Air Susu Ibu (ASI)
Menggunakan *Multivariate Adaptive Regression Splines* (MARS) dan
Algoritma C4.5 di Kabupaten Sragen**

Disusun Oleh:

YUSUF ARIFKA RAHMAN

24010211120011

Skripsi

Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Sains pada
Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro

JURUSAN STATISTIKA

FAKULTAS SAINS DAN MATEMATIKA

UNIVERSITAS DIPONEGORO

SEMARANG

2015

HALAMAN PENGESAHAN I

Judul : Ketepatan Klasifikasi Status Pemberian Air Susu Ibu (ASI)
Menggunakan *Multivariate Adaptive Regression Splines* (MARS)
dan Algoritma C4.5 di Kabupaten Sragen
Nama : Yusuf Arifka Rahman
NIM : 24010211120011

Telah diujikan pada sidang Tugas Akhir tanggal 22 Desember 2015 dan
dinyatakan lulus pada tanggal 30 Desember 2015

Semarang, Desember 2015

Mengetahui,

Ketua Jurusan Statistika
FSM UNDIP



Dra. Hj. Dwi Ispriyanti, M.Si

NIP. 195709141986032001

Panitia Penguji Ujian Tugas Akhir
Ketua,

Drs. Tarno, M.Si

NIP. 196307061991021001

HALAMAN PENGESAHAN II

Judul : Ketepatan Klasifikasi Status Pemberian Air Susu Ibu (ASI)
Menggunakan *Multivariate Adaptive Regression Splines* (MARS)
dan Algoritma C4.5 di Kabupaten Sragen
Nama : Yusuf Arifka Rahman
NIM : 24010211120011

Telah diujikan pada sidang Tugas Akhir tanggal 22 Desember 2015.

Semarang, Desember 2015

Dosen Pembimbing I



Dra. Suparti, M.Si

NIP. 196509131990032001

Dosen Pembimbing II



Sugito, S.Si, M.Si

NIP. 197610192005011001

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT atas rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang diberi judul **“Ketepatan Klasifikasi Status Pemberian Air Susu Ibu (ASI) Menggunakan *Multivariate Adaptive Regression Splines (MARS)* dan Algoritma C4.5 di Kabupaten Sragen”**. Tugas Akhir ini tidak akan terselesaikan dengan baik tanpa adanya dukungan dan bantuan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada:

1. Ibu Dra. Hj. Dwi Ispriyanti, M.Si. selaku Ketua Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro.
2. Ibu Dra. Suparti, M.Si selaku dosen pembimbing I
3. Bapak Sugito, S.Si, M.Si selaku dosen pembimbing II.
4. Bapak/Ibu dosen jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro.
5. Semua pihak yang telah membantu kelancaran penyusunan Tugas Akhir ini, yang tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa dalam penyusunan Tugas Akhir ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran demi kesempurnaan penulisan selanjutnya.

Semarang, Desember 2015

Penulis

ABSTRAK

Kemajuan suatu bangsa dipengaruhi dan ditentukan oleh tingkat kesehatan masyarakat di mana salah satu indikator tingkat kesehatan tersebut ditentukan oleh status gizi. Gizi dapat diberikan sejak dini yaitu dengan pemberian Air Susu Ibu (ASI) kepada bayi. Penelitian ini bertujuan untuk membandingkan pengklasifikasian pemberian ASI secara eksklusif dan tidak eksklusif. Digunakan dua metode untuk melakukan klasifikasi pemberian ASI kepada bayi di Kabupaten Sragen tahun 2014 yaitu *Multivariate Adaptive Regression Splines* (MARS) dan algoritma C4.5. Metode MARS merupakan metode regresi nonparametrik yang digunakan untuk mengatasi data berdimensi tinggi yang menghasilkan prediksi akurat serta model yang kontinu dalam knot. Algoritma C4.5 adalah salah satu metode klasifikasi dari *data mining* yang digunakan untuk mengkonstruksikan pohon keputusan (*decision tree*). Untuk mengevaluasi hasil klasifikasi digunakan perhitungan *Apparent Error Rate* (APER). Hasil klasifikasi terbaik menggunakan metode MARS adalah menggunakan kombinasi *Basis Function* (BF)=40, *Maximum Interaction* (MI)=3, *Minimum Observation* (MO)=3 karena menghasilkan nilai *Generalized Cross Validation* (GCV) terkecil. Hasil klasifikasi metode mars diperoleh nilai APER sebesar 19,7674% dan ketepatan klasifikasi sebesar 80,2326%. Hasil klasifikasi algoritma C4.5 diperoleh nilai APER sebesar 18,6047% dan ketepatan klasifikasi sebesar 81,3953%. Dengan melakukan uji proporsi, disimpulkan kedua metode tersebut menghasilkan klasifikasi yang sama baiknya.

Kata kunci: Pemberian ASI, Klasifikasi, MARS, Algoritma C4.5

ABSTRACT

The progress of a nation influenced and determined by the level of public health, the indicator of the level of health is determined by nutritional status. Nutrition can be given early, namely breastfeeding to infants. This research aims to compare the classification of exclusive breastfeeding and nonexclusive breastfeeding. It used two methods for classifying a breastfeeding to babies in Sragen subdistrict on 2014, the methods are Multivariate Adaptive Regression Splines (MARS) and C4.5 Algorithm. MARS is nonparametric regression method that use to overcome the high dimension of data that produces accurate prediction and continuous models on knot. C4.5 Algorithm is a way of classifying methods from data mining that use to construct a decision tree. To evaluate the result of classification use Apparent Error Rate (APER) calculation. The best classification result using MARS method is by using the combination of Basis Function (BF)=40, Maximum Interaction (MI)=3, Minimum Observation (MO)=3 because it will result on the smallest Generalized Cross Validation (GCV). Classification result using MARS method obtained APER is 19,7674% and 80,2326% of accuracy. Classification result using C4.5 Algorithm obtained APER is 18,6047% and 81,3953% of accuracy. From proportion test, concluded classification that formed by MARS is as good as by C4.5 Algorithm.

Keywords: Breastfeeding, Classification, MARS, C4.5 Algorithm

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
LEMBAR PENGESAHAN I.....	ii
LEMBAR PENGESAHAN II.....	iii
KATA PENGANTAR.....	iv
ABSTRAK.....	v
ABSTRACT.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN.....	xii
BAB I PENDAHULUAN	
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	5
1.3. Batasan Masalah.....	5
1.4. Tujuan Penelitian.....	6
BAB II TINJAUAN PUSTAKA	
2.1. Air Susu Ibu (ASI).....	7
2.2. <i>Multivariate Adaptive Regression Spline</i> (MARS).....	9
2.2.1. Estimasi Parameter.....	12
2.2.2. Anova Dekomposisi.....	15
2.2.3. Klasifikasi MARS respon biner.....	18

2.3. Data Mining.....	19
2.4. Algoritma C4.5	21
2.5. Ketepatan Klasifikasi.....	26
2.6. Uji Hipotesis Dua Proporsi.....	28
BAB III METODOLOGI PENELITIAN	
3.1. Jenis dan Sumber Data	29
3.2. Variabel Penelitian	29
3.3. Langkah-langkah Analisis.....	31
3.4. Diagram Alir (<i>Flowchart</i>)	33
BAB IV HASIL DAN PEMBAHASAN	
4.1. Analisis <i>Multivariate Adaptive Regression Splines</i> (MARS)	34
4.1.1. Pemodelan MARS	34
4.1.2. Model Terbaik	36
4.1.3. Interpretasi Model MARS	37
4.1.4. Klasifikasi Model MARS	39
4.2. Pengklasifikasian Menggunakan Algoritma C4.5	42
4.2.1. Konstruksi Algoritma C4.5	42
4.2.2. Pemangkasan Pohon Klasifikasi.....	48
4.2.3. Analisis Pohon Keputusan.....	52
4.2.4. Identifikasi Pemberian ASI	54
4.2.5. Klasifikasi Algoritma C4.5.....	55
4.3. Ringkasan Ketepatan Klasifikasi Kedua Metode	56
4.4. Evaluasi Ketepatan Klasifikasi.....	57
BAB V KESIMPULAN	59

DAFTAR PUSTAKA 61

LAMPIRAN 63



DAFTAR TABEL

	Halaman
Tabel 1 Matriks Konfusi untuk Klasifikasi Dua Kelas.....	27
Tabel 2 Nilai GCV dari Beberapa Kombinasi BF, MI, dan MO	34
Tabel 3 Klasifikasi Data <i>Testing</i> Metode MARS.....	41
Tabel 4 Frekuensi Tiap Kelas	42
Tabel 5 Peluang Masing-masing Kelas	43
Tabel 6 Nilai Ambang Batas Atribut Umur ibu pada Pohon Keputusan.....	44
Tabel 7 Frekuensi Masing-masing Kategori pada Atribut Umur Ibu Berdasarkan Kelasnya	45
Tabel 8 Peluang Masing-masing Kategori Berdasarkan Kelasnya.....	45
Tabel 9 Nilai <i>Information Gain</i> pada Simpul Akar	46
Tabel 10 Klasifikasi Data <i>Training</i> Algoritma C4.5	54
Tabel 11 Klasifikasi Data <i>Testing</i> Algoritma C4.5	56
Tabel 12 Nilai Ketepatan Klasifikasi Data <i>Testing</i> Dua Metode	57

DAFTAR GAMBAR

	Halaman
Gambar 1 Contoh Pohon Keputusan (<i>Decision Tree</i>)	21
Gambar 2 Diagram Alir Analisis Data	33
Gambar 3 Pohon Keputusan Tingkat Pertama	47
Gambar 4 Pohon Klasifikasi tanpa Pemangkasan	48
Gambar 5 Simpul Keputusan Ijazah Terakhir Kepala Rumah Tangga	49
Gambar 6 Pohon Keputusan dengan Pemangkasan	52

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Data SUSENAS Kabupaten Sragen Tahun 2014	63
Lampiran 2 Masing-masing Model dengan Nilai GCV Terkecil.....	64
Lampiran 3 Analisis <i>Multivariate Adaptive Regression Spline</i> (MARS)	66
Lampiran 4 Hasil Algoritma C4.5 Menggunakan Data <i>Training</i>	70
Lampiran 5 Nilai gain dari semua atribut untuk menentukan simpul akar	72
Lampiran 6 Pohon Keputusan Sebelum Pemangkasan.....	76
Lampiran 7 Pohon Keputusan Setelah Pemangkasan.....	77
Lampiran 8 Hasil Algoritma C4.5 Menggunakan Data <i>Testing</i>	78
Lampiran 9 Hasil Prediksi Data <i>Testing</i> Algoritma C4.5	79
Lampiran 10 Hasil Prediksi Data <i>Testing</i> Kombinasi BF=40 MI=3 dan MO=3 Metode MARS	83

BAB I

PENDAHULUAN

1.1 Latar Belakang

Air Susu Ibu (ASI) adalah makanan terbaik bagi bayi karena mengandung zat gizi paling sesuai untuk pertumbuhan dan perkembangan bayi (Soetjiningsih, 1997). Pemberian ASI akan menjamin bayi tetap sehat dan memulai kehidupan dengan cara yang paling sehat. Dengan pemberian ASI tidak saja memberikan kesempatan pada bayi untuk tumbuh menjadi manusia yang sehat secara fisik, tetapi juga lebih cerdas, mempunyai emosional yang lebih stabil, perkembangan spiritual yang positif, serta perkembangan sosial yang lebih baik (Roesli, 2005).

Berdasarkan Keputusan Menteri Kesehatan Republik Indonesia Nomor 450 Tahun 2004 tentang Air Susu Ibu (ASI) secara eksklusif bagi bayi di Indonesia yaitu diberikan sejak bayi lahir sampai dengan bayi berumur 6 (enam) bulan dan dianjurkan dilanjutkan sampai anak berusia 2 (dua) tahun dengan pemberian makanan tambahan yang sesuai. Hak bayi dalam kesempatannya untuk mendapatkan ASI tercantum dalam UU No. 36 pasal 128 ayat 1 yang berisi bahwa setiap bayi berhak mendapatkan air susu ibu eksklusif sejak dilahirkan selama 6 (enam) bulan, kecuali atas indikasi medis (Asosiasi Ibu Menyusui Indonesia, 2012). Menurut Diharjo dalam Haryani (2014), ibu yang menyusui secara eksklusif mempunyai kontribusi yang cukup besar terhadap peningkatan derajat kesehatan bayi terutama menurunnya jumlah kematian bayi. Oleh karena itu sangat disayangkan apabila sesudah persalinan

ibu tidak memberikan ASI secara eksklusif atau bahkan menghentikan sama sekali pemberian ASI kepada bayinya.

Beberapa faktor diduga menyebabkan bayi tidak mendapatkan ASI dengan baik, faktor tersebut adalah faktor karakteristik ibu, faktor bayi, lingkungan, dukungan keluarga, pendidikan kesehatan, sosial ekonomi dan budaya (Budiharjo, 2003). Selain itu berdasarkan beberapa laporan studi tentang permasalahan pemberian ASI eksklusif menemukan faktor-faktor tidak diberikannya ASI eksklusif pada bayi adalah karena ibu sibuk bekerja, pendidikan ibu yang rendah, gencarnya periklanan tentang penggunaan susu formula, kurangnya sekresi ASI, persepsi tentang bayi tanpa diberi makanan tambahan akan menjadi lapar dan pengetahuan ibu tentang ASI kurang (Diharjo dalam Haryani, 2014). Tenaga kerja perempuan yang meningkat menjadi salah satu kendala dalam mensukseskan program ASI Eksklusif, hal ini karena cuti melahirkan hanya 12 minggu, dimana 4 (empat) minggu diantaranya sering harus diambil sebelum melahirkan (Suradi, 2003). Dengan demikian, ibu yang bekerja hanya dapat mendampingi bayinya secara intensif hanya 2 (dua) bulan, termasuk dalam penyusuan bayinya. Setelah itu ibu harus kembali bekerja dan sering kali ibu terpaksa berhenti menyusui.

Menurut Siregar (2004), beberapa faktor yang mempengaruhi pemberian ASI eksklusif pada bayi antara lain adalah faktor pendidikan, pengetahuan, pekerjaan, ekonomi, budaya, psikologis, informasi susu formula dan kesehatan. Seseorang berpendidikan tinggi dan berpengetahuan luas akan lebih bisa menerima alasan untuk memberikan ASI eksklusif karena pola pikirnya yang lebih realistis dibandingkan

yang tingkat pendidikan rendah. Kesibukan sosial lain serta kenaikan tingkat partisipasi wanita dalam angkatan kerja dan adanya emansipasi dalam segala bidang kerja dan di kebutuhan masyarakat menyebabkan turunnya kesediaan menyusui dan lamanya menyusui. Selain itu faktor yang menjadi penyebab kurangnya pemberian ASI eksklusif yaitu belum semua petugas paramedis diberi pesan dan diberi cukup informasi agar menganjurkan setiap ibu untuk menyusui bayi mereka, serta praktek yang keliru dengan memberikan susu botol kepada bayi yang baru lahir. Selain itu kemampuan ibu yang usianya lebih tua juga amat rendah produksi ASInya, sehingga bayi cenderung mengalami malnutrisi.

Pada tugas akhir ini digunakan dua metode untuk mengklasifikasikan data pemberian ASI eksklusif tersebut, yaitu *Multivariate Adaptive Regression Splines* (MARS) dan Algoritma C4.5. Metode *Multivariate Adaptive Regression Splines* (MARS) merupakan pendekatan regresi non parametrik yang dikembangkan oleh Jerome H. Friedman pada tahun 1990. Untuk selanjutnya, metode ini disingkat dengan MARS. MARS adalah sebuah metode dalam regresi multivariat yang membantu dalam memodelkan hubungan antara suatu variabel respon dengan beberapa variabel prediktor. MARS merupakan salah satu pendekatan non parametrik sehingga tidak tergantung pada asumsi bentuk model tertentu. Menurut Friedman (1991), model MARS berguna untuk mengatasi data berdimensi tinggi dan dapat menghasilkan prediksi variabel respon yang akurat.

Algoritma C4.5 adalah salah satu metode klasifikasi dari *data mining* yang digunakan untuk mengkonstruksikan pohon keputusan (*decision tree*). Algoritma

C4.5 diperkenalkan oleh Quinlan (1993) sebagai versi perbaikan dari algoritma *Iterative Dichotomiser 3* (ID3). Secara umum, terdapat enam fungsi dalam *data mining* untuk menemukan atau menggali pengetahuan dalam data yang berukuran besar, yaitu fungsi deskripsi, estimasi, prediksi, klasifikasi, *clustering*, dan asosiasi. Ketiga fungsi pertama dikenal sebagai fungsi minor (fungsi tambahan) dan ketiga fungsi terakhir dikenal sebagai fungsi mayor (fungsi utama). Menurut Wu, et al. (2007), terdapat sejumlah algoritma fungsi *data mining* yang dipilih menjadi *Top 10 algorithm data mining*. Hasilnya, diperoleh bahwa algoritma klasifikasi menjadi mayoritas dalam *Top 10 algorithm*. Adapun algoritma klasifikasi yang terbaik dalam *Top 10 algorithm data mining* adalah pohon keputusan C4.5, *k-Nearest Neighbor (k-NN)*, *Naive Baiyes*, dan *Classification and Regression Trees (CART)*.

Sehubungan dengan tugas akhir ini, akan diteliti mengenai data status pemberian ASI di Kabupaten Sragen menggunakan metode *Multivariate Adaptive Regression Splines* (MARS) dan Algoritma C4.5. Dari kedua metode tersebut akan dicari nilai ketepatan klasifikasi data *testing* yang terbaik. Sehingga pada penelitian ini penulis memilih judul “Ketepatan Klasifikasi Status Pemberian Air Susu Ibu (ASI) Menggunakan *Multivariate Adaptive Regression Splines* (MARS) dan Algoritma C4.5 di Kabupaten Sragen”.

1.2 Perumusan Masalah

Berdasarkan latar belakang, permasalahan yang akan dibahas pada penelitian ini adalah:

1. Bagaimana ketepatan klasifikasi status pemberian ASI di kabupaten Sragen menggunakan metode *Multivariate Adaptive Regression Splines* (MARS)?
2. Bagaimana penerapan Algoritma C4.5 dalam pembentukan pohon klasifikasi untuk data status pemberian ASI di kabupaten Sragen?
3. Bagaimana ketepatan klasifikasi status pemberian ASI di kabupaten Sragen menggunakan metode Algoritma C4.5?
4. Bagaimana perbandingan ketepatan klasifikasi berdasarkan analisis *Multivariate Adaptive Regression Splines* (MARS) dan algoritma C4.5?

1.3 Pembatasan Masalah

Dalam penelitian ini permasalahan dibatasi pada perbandingan hasil ketepatan klasifikasi menggunakan metode *Multivariate Adaptive Regression Splines* (MARS) dan Algoritma C4.5. Data yang digunakan dalam penelitian ini adalah data hasil Survei Sosial Ekonomi Nasional (SUSENAS) di Kabupaten Sragen tahun 2014 yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Jawa Tengah.

1.4 Tujuan Penelitian

Berdasarkan permasalahan yang telah disampaikan tujuan penelitian ini adalah:

1. Menentukan ketepatan klasifikasi status pemberian ASI di kabupaten Sragen berdasarkan analisis *Multivariate Adaptive Regression Splines* (MARS).
2. Membentuk pohon klasifikasi untuk data status pemberian ASI di kabupaten Sragen menggunakan Algoritma C4.5.
3. Menentukan ketepatan klasifikasi status pemberian ASI di kabupaten Sragen berdasarkan analisis algoritma C4.5.
4. Membandingkan ketepatan klasifikasi berdasarkan analisis *Multivariate Adaptive Regression Splines* (MARS) dan algoritma C4.5.