

**PERBANDINGAN METODE REGRESI LOGISTIK BINER
DAN *MULTIVARIATE ADAPTIVE REGRESSION SPLINE***

(MARS) PADA PEMINATAN JURUSAN SMA

(Studi Kasus SMA Negeri 2 Semarang)



=====
SKRIPSI
=====

Oleh :

RATIH BINADARI

24010211120022

JURUSAN STATISTIKA

FAKULTAS SAINS DAN MATEMATIKA

UNIVERSITAS DIPONEGORO

SEMARANG

2015

**PERBANDINGAN METODE REGRESI LOGISTIK BINER
DAN *MULTIVARIATE ADAPTIVE REGRESSION SPLINE*
(MARS) PADA PEMINATAN JURUSAN SMA
(Studi Kasus SMA Negeri 2 Semarang)**

Oleh :

RATIH BINADARI

24010211120022

Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Sains pada Jurusan Statistika

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

2015

HALAMAN PENGESAHAN I

Judul Skripsi : Perbandingan Metode Regresi Logistik Biner dan *Multivariate Adaptive Regression Spline* (MARS) terhadap Peminatan Jurusan SMA (Studi kasus SMA Negeri 2 Semarang)

Nama : Ratih Binadari

NIM : 24010211120022

Jurusan : Statistika

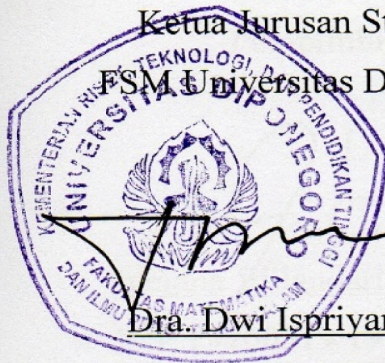
Telah diujikan pada sidang Tugas Akhir tanggal 11 September 2015 dan dinyatakan lulus pada tanggal 29 September 2015.

Semarang, 30 September 2015

Mengetahui,

Ketua Jurusan Statistika

FSM Universitas Diponegoro



Dra. Dwi Ispriyanti, M.Si.

NIP. 19751008 200312 2 001

Panitia Penguji Ujian Tugas Akhir

Ketua,

Drs. Agus Rusgiyono, M.Si.

NIP. 19640813 199001 1 001

HALAMAN PENGESAHAN II

Judul Skripsi : Perbandingan Metode Regresi Logistik Biner dan *Multivariate Adaptive Regression Spline* (MARS) terhadap Peminatan Jurusan SMA (Studi kasus SMA Negeri 2 Semarang)

Nama : Ratih Binadari

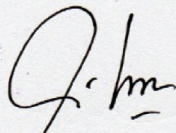
NIM : 24010211120022

Jurusan : Statistika

Telah diujikan pada sidang Tugas Akhir tanggal 11 September 2015.

Semarang, 30 September 2015

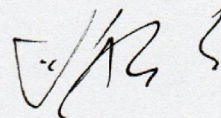
Pembimbing I



Yuciana Wilandari, S.Si, M.Si.

NIP 19700519 199802 2 001

Pembimbing II



Dra. Suparti, M.Si.

NIP. 19650913 199003 2 001

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas segala limpahan rahmat, dan hidayah-Nya sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul “Perbandingan Metode Regresi Logistik Biner dan *Multivariate Adaptive Regression Spline* (MARS) pada Peminatan Jurusan SMA (Studi Kasus SMA Negeri 2 Semarang)”

Penulisan Tugas Akhir ini tidak akan selesai dengan baik tanpa adanya dukungan dan bantuan dari berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada :

1. Ibu Dra. Dwi Ispriyanti, M.Si. selaku Ketua Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro.
2. Bapak Moch. Abdul Mukid, S.Si, M.Si, selaku Koordinator Seminar dan Tugas Akhir.
3. Ibu Yuciana Wilandari, S.Si, M.Si, selaku dosen pembimbing I dan Ibu Suparti, S.Si, M.Si, selaku dosen pembimbing II seminar dan tugas akhir.
4. Semua pihak yang terkait dalam penulisan tugas akhir ini.

Penulis menyadari bahwa laporan ini masih banyak kekurangan, baik pada teknis penulisan maupun materi yang disampaikan. Oleh karena itu, kritik dan saran dari semua pihak yang bersifat membangun penulis harapkan demi kesempurnaan laporan ini. Semoga Tugas Akhir ini bermanfaat bagi banyak pihak.

Semarang, September 2015

Penulis

ABSTRAK

Peminatan jurusan SMA bertujuan untuk memberikan peluang yang terbuka pada peserta didik untuk memilih mata pelajaran yang diminati dan mengembangkan berbagai potensi yang dimilikinya sesuai dengan kemampuan, minat, bakat, dan kepribadian. Dalam peminatan jurusan SMA, ada beberapa faktor yang mempengaruhinya. Untuk mengetahui faktor-faktor tersebut, maka digunakan metode regresi logistik biner dan *Multivariate Adaptive Regression Spline* (MARS). Metode regresi logistik biner adalah metode yang menjelaskan hubungan antara variabel respon dan beberapa variabel prediktor dengan variabelnya bernilai 1 untuk menyatakan keberadaan suatu karakteristik dan bernilai 0 untuk menyatakan ketidakberadaan suatu karakteristik. Metode MARS merupakan metode regresi nonparametrik multivariat pengembangan metode *Recursive Partitioning Regression* (RPR) dan *spline* untuk data berdimensi tinggi yang menghasilkan prediksi akurat serta model yang kontinu dalam knot. Kedua metode tersebut dibandingkan guna mengetahui metode terbaik yang digunakan dalam penelitian. Dari hasil analisis, baik menggunakan metode regresi logistik biner ataupun metode MARS, disimpulkan bahwa faktor yang mempengaruhi peminatan jurusan SMA adalah Nilai UN Matematika, Nilai UN IPA dan Hubungan dengan Teman. Dengan melakukan uji proporsi, disimpulkan bahwa kedua metode tersebut menghasilkan klasifikasi yang sama baiknya.

Kata kunci: Peminatan Jurusan SMA, Regresi Logistik, *Multivariate Adaptive Regression Spline* (MARS), klasifikasi

ABSTRACT

Major specialization at High School is aimed to gives opened opportunity for students to choose subject that are interest and develop their potential in accordance with the abilities, interests, talents, and personality. Major specialization at High school is influenced by some factors. To detect those factors, used biner logistic regression method and Multivariate Adaptive Regression Spline (MARS). Biner Logistic Regression is method that describes relationship between dependent variable and some independent variable, with independent variable has been coded 1 as representing the presence of the characteristic, and 0 as representing the absence of the characteristic. MARS is multivariate nonparametric regression method that development of Recursive Partitioning Regression (RPR) method and Spline method for high dimensional data that produces accurate prediction and continuous models on knots. Both of the methods are compared to know the best method used in research. From the result of analysis using biner logistic regression method and MARS, concluded that major specialization has been influenced by mathematic score, science score and relationship between students and friends. From proportion test, concluded classification that formed by regression logistic is as good as by MARS.

Keywords : Major specialization at High School, Biner Logistic Regression, Mutlivariate Adaptive Regression Spline (MARS), Clasification

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
KATA PENGANTAR	iv
ABSTRAK	v
ABSTRACT.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR LAMPIRAN.....	xi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Perumusan Masalah.....	4
1.3. Pembatasan Masalah.....	5
1.4. Tujuan.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1. Pendidikan di Indonesia	6
2.2. Peminatan Jurusan SMA	7
2.3. Uji Validitas dan Reliabilitas.....	8
2.4. Analisis Regresi Logistik	10
2.4.1. Metode Regresi Logistik Biner.....	11
2.4.2. Estimasi Parameter	12
2.4.3. Uji Signifikansi Parameter.....	16
2.4.4. Uji <i>Goodness of Fit</i>	18

2.4.5. Klasifikasi pada Regresi Logistik.....	19
2.5. <i>Multivariate Adaptive Regression Spline</i> (MARS)	19
2.5.1. Model MARS	21
2.5.2. Estimasi Model MARS.....	23
2.5.3. Dekomposisi ANOVA	26
2.5.4. Pemilihan Model Terbaik	27
2.5.5. Klasifikasi pada MARS	28
2.6. Evaluasi Ketepatan Klasifikasi.....	29
BAB III METODOLOGI PENELITIAN	31
3.1. Jenis dan Sumber Data	31
3.2. Metode Pengumpulan Data	31
3.3. Populasi dan Sampel.....	31
3.4. Variabel Penelitian	32
3.5. Langkah-Langkah Analisis.....	33
3.6. Diagram Alur (<i>Flowchart</i>).....	35
BAB IV HASIL DAN PEMBAHASAN	37
4.1. Uji Validitas dan Reliabilitas.....	37
4.2. Analisis Regresi Logistik Biner.....	40
4.2.1. Estimasi Parameter	40
4.2.2. Uji Signifikansi Parameter.....	41
4.2.3. Uji <i>Goodness of Fit</i>	44
4.2.4. Model Akhir	45
4.2.5. Klasifikasi	46
4.3. Analisis MARS	48

4.3.1. Model Terbaik.....	49
4.3.2. Interpretasi Model MARS	50
4.3.3. Klasifikasi	51
4.4. Evaluasi Ketepatan Klasifikasi.....	53
BAB V KESIMPULAN	59
DAFTAR PUSTAKA	61
LAMPIRAN	

DAFTAR TABEL

	Halaman
Tabel 1. Matriks Konfusi Klasifikasi Dua Kelas	29
Tabel 2. Distribusi Sampel Penelitian.....	32
Tabel 3. Variabel Penelitian.....	33
Tabel 4. Hasil Uji Validitas Variabel Hubungan Siswa dengan Teman	37
Tabel 5. Hasil Uji Validitas Variabel Hubungan Siswa dengan Guru.....	38
Tabel 6. Hasil Uji Validitas Variabel Hubungan Siswa dengan Keluarga	38
Tabel 7. Hasil Uji Validitas Variabel Motivasi dari Diri Sendiri	38
Tabel 8. Hasil Uji Reliabilitas.....	39
Tabel 9. Estimasi Parameter untuk Model Awal Regresi Logistik	40
Tabel 10. Nilai Wald untuk Setiap Parameter.....	42
Tabel 11. Nilai Wald untuk X_1 , X_2 , X_{4_1} , dan X_{4_2}	44
Tabel 12. Hasil Klasifikasi Metode Regresi Logistik Biner	47
Tabel 13. <i>Trial and Error</i> Kombinasi BF, MI dan MO.....	48
Tabel 14. Hasil Klasifikasi Metode MARS	52

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Kuisisioner Penelitian	59
Lampiran 2. Data Uji Validitas dan Uji Reliabilitas	63
Lampiran 3. Data Penelitian.....	65
Lampiran 4. Uji Validitas.....	67
Lampiran 5. Uji Reliabilitas	69
Lampiran 6. Analisis Regresi Logistik Biner terhadap Peminatan Jurusan SMA	71
Lampiran 7. Analisis MARS terhadap Peminatan Jurusan SMA	77
Lampiran 8. Tabel Distribusi <i>Chi Square</i> (χ^2)	80
Lampiran 9. Tabel r	81
Lampiran 10. Tabel Distribusi Normal Kumulatif Z	82
Lampiran 11. Mengkategorikan Data	84

BAB I

PENDAHULUAN

1.1. Latar Belakang

Pendidikan adalah usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta keterampilan yang diperlukan dirinya, masyarakat, bangsa dan negara (Undang-Undang Republik Indonesia Nomor 20 Tahun 2003, Pasal 1 Ayat 1). Di Indonesia, pendidikan dapat dibedakan berdasarkan jalur, jenjang dan jenisnya. Jalur pendidikan dibagi menjadi tiga yaitu jalur formal atau pendidikan yang diselenggarakan di sekolah-sekolah pada umumnya, jalur nonformal atau pendidikan yang didapat dari luar seperti kursus dan pendidikan anak usia dini, serta jalur informal atau jalur pendidikan keluarga dan lingkungan berbentuk kegiatan belajar secara mandiri. Jenjang pendidikan jalur formal ada tiga yaitu pendidikan dasar, pendidikan menengah dan pendidikan tinggi. Sedangkan jika dilihat dari jenisnya, jenis pendidikan mencakup pendidikan umum, kejuruan, akademik, profesi, vokasi, keagamaan, dan khusus. Hal ini dijelaskan di dalam Undang-Undang Republik Indonesia Nomor 20 Tahun 2003.

Sekolah Menengah Atas (SMA) merupakan pendidikan umum jalur formal jenjang sekolah menengah yang biasanya ditempuh dalam waktu 3 tahun. Pada jenjang Sekolah Menengah Atas, peserta didik akan mengikuti peminatan jurusan. Menurut Sudiarto (2013), tujuan diadakannya peminatan jurusan adalah untuk

memberikan peluang yang lebih terbuka kepada peserta didik untuk memilih mata pelajaran yang diminati, mendalami materi mata pelajaran dan mengembangkan berbagai potensi yang dimilikinya secara fleksibel sesuai dengan kemampuan dasar umum (kecerdasan), bakat, minat dan karakteristik kepribadian tanpa dibatasi dengan sekat-sekat penjurusan yang terlalu kaku. Peminatan jurusan di SMA sangat penting karena dapat menjadi landasan dan acuan ke depan dalam meraih cita-cita serta dalam memilih jurusan di bangku perkuliahan. Umumnya pilihan minat jurusan di SMA ada tiga yaitu Ilmu Alam, Ilmu Sosial dan Ilmu Bahasa.

Ada beberapa faktor yang mempengaruhi peserta didik dalam peminatan jurusan. Menurut Syah (2009), faktor yang mempengaruhi antara lain faktor internal dan faktor eksternal. Faktor internal yaitu faktor yang datang dari dalam diri siswa itu sendiri, sedangkan faktor eksternal adalah faktor yang datang dari luar diri siswa tersebut. Untuk mengetahui faktor apa saja yang mempengaruhi peserta didik dalam peminatan jurusan, maka perlu diadakan sebuah penelitian.

SMA Negeri 2 Semarang adalah satu dari 16 SMA Negeri di Semarang yang terletak di Kecamatan Pedurungan. Sebelum tahun 2010, minat jurusan yang disediakan SMA Negeri 2 Semarang adalah Ilmu Alam, Ilmu Sosial dan Ilmu Bahasa. Namun setelah tahun 2010 pilihan minat jurusan Ilmu Bahasa ditiadakan karena kurangnya minat dari peserta didik. Jumlah peserta didik SMA Negeri 2 Semarang Tahun Ajaran 2012/2013 hingga Tahun Ajaran 2014/2015 adalah sebanyak 1.438 siswa dengan jumlah peserta didik kelas X adalah 505 siswa, kelas XI ada 470 siswa dan kelas XII berjumlah 463 siswa.

Dalam ilmu statistika, banyak metode yang dapat digunakan untuk mengetahui pengaruh variabel prediktor terhadap variabel responnya yang bersifat kategori. Diantaranya adalah metode regresi logistik biner dan *Multivariate Adaptive Regression Spline* (MARS). Kedua metode tersebut termasuk dalam analisis regresi. Menurut Montgomery dan Peck (1992) analisis regresi adalah suatu analisis statistika yang memanfaatkan hubungan antara dua variabel atau lebih. Umumnya analisis ini digunakan untuk melihat hubungan dan pengaruh variabel prediktor terhadap variabel respon dimana variabel responnya berupa data kuantitatif. Namun untuk variabel respon yang bersifat kualitatif/kategori dapat dianalisis menggunakan metode regresi logistik. Hosmer dan Lemeshow (2000) mengatakan bahwa metode regresi logistik adalah suatu metode analisis statistika yang mendeskripsikan hubungan antara variabel respon yang memiliki dua kategori atau lebih dengan satu atau lebih variabel prediktor. Salah satu model regresi logistik adalah regresi logistik biner. Model regresi logistik biner merupakan metode regresi logistik yang digunakan untuk menganalisis hubungan antara satu variabel respon dan beberapa variabel prediktor, dengan variabel responnya berupa data kualitatif dikotomi yaitu bernilai 1 untuk menyatakan keberadaan sebuah karakteristik dan bernilai 0 untuk menyatakan ketidakberadaan sebuah karakteristik.

Metode statistika lainnya yang dapat digunakan juga untuk pengklasifikasian objek dengan melibatkan variabel respon dan kategori dengan sejumlah variabel prediktor kontinu ataupun kategori adalah metode *Multivariate Adaptive Regression Spline* (MARS). MARS merupakan salah satu model regresi nonparametrik multivariat yang dikembangkan oleh Friedman tahun 1991.

Metode ini merupakan pengembangan dari pendekatan *Recursive Partitioning Regression* (RPR) dengan metode *Spline* (Friedman, 1991). Model MARS berguna untuk mengatasi permasalahan data berdimensi tinggi dan menghasilkan prediksi variabel respon yang akurat serta menghasilkan model kontinu dalam *knot* berdasarkan nilai *Generalized Cross Validation* (GCV) terkecil. Perbandingan metode regresi logistik dan MARS juga sudah pernah dilakukan sebelumnya pada penelitian yang berjudul “*Ketepatan Klasifikasi Dengan Analisis Regresi Logistik dan Multivariate Adaptive Regression Spline (MARS) pada Data dengan Peubah Data Biner*” (Fitrianty et al. 2013) dan “*Pemodelan terhadap Kelulusan Siswa Masuk Akselerasi Menggunakan Analisis Regresi Logistik dan Multivariate Adaptive Regression Spline (MARS)*” (Mandaku dan Mandaku, 2010).

Sehubungan dengan penelitian ini, maka variabel yang diduga mempengaruhi peserta didik dalam peminatan jurusan terdiri dari faktor internal yang meliputi nilai UN Matematika, nilai UN IPA, nilai UN Bahasa Inggris dan dorongan dari diri sendiri, serta faktor eksternal yang meliputi hubungan siswa dengan teman, hubungan siswa dengan guru, dan hubungan siswa dengan keluarga. Dengan demikian, penelitian ini dibuat dengan judul “*Perbandingan Metode Regresi Logistik Biner dan Multivariate Adaptive Regression Spline (MARS) pada Peminatan Jurusan SMA (Studi Kasus SMA Negeri 2 Semarang)*”

1.2. Perumusan Masalah

Permasalahan yang dibahas pada penelitian ini adalah menganalisis serta memodelkan faktor-faktor yang mempengaruhi peminatan jurusan SMA dengan

menggunakan dua metode yaitu metode Regresi Logistik Biner dan metode MARS yang kemudian kedua metode tersebut dibandingkan untuk mendapatkan metode terbaik dalam pemodelan dan klasifikasi peminatan jurusan SMA.

1.3. Pembatasan Masalah

Data yang digunakan dalam penelitian ini adalah data mengenai peminatan jurusan siswa kelas X, XI dan XII di SMA Negeri 2 Semarang tahun ajaran 2014/2015. Metode yang digunakan dalam penelitian ini adalah metode Regresi Logistik Biner dan *Multivariate Adaptive Regression Spline* (MARS).

1.4. Tujuan

Tujuan yang ingin didapat dari penelitian ini adalah:

1. Menentukan faktor-faktor yang mempengaruhi peserta didik dalam peminatan jurusan di SMA Negeri 2 Semarang.
2. Mendapatkan pemodelan peminatan jurusan SMA dari metode Regresi Logistik Biner dan MARS
3. Membandingkan metode Regresi Logistik Biner dan MARS untuk mendapatkan metode terbaik dalam pemodelan dan klasifikasi peminatan jurusan di SMA Negeri 2 Semarang.