

**PENDUGAAN PARAMETER REGRESI LOGISTIK
MULTINOMIAL MENGGUNAKAN METODE *LEAST ABSOLUTE
SHRINKAGE AND SELECTION OPERATOR* (LASSO)
(Studi Kasus Pilihan Sekolah SMA Sederajat Oleh Siswa SMP dan
MTs Negeri di Kabupaten Trenggalek)**

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains
dalam bidang Statistika

Oleh:

HAFID WAHYU RAMADHAN

135090500111008



**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2017**

**PENDUGAAN PARAMETER REGRESI LOGISTIK
MULTINOMIAL MENGGUNAKAN METODE *LEAST
ABSOLUTE SHRINKAGE AND SELECTION OPERATOR*
(LASSO)**

**(Studi Kasus Pilihan Sekolah SMA Sederajat Oleh Siswa SMP
dan MTs Negeri di Kabupaten Trenggalek)**

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains
dalam bidang Statistika

Oleh:
HAFID WAHYU RAMADHAN
135090500111008



**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2017**

LEMBAR PENGESAHAN SKRIPSI

**PENDUGAAN PARAMETER REGRESI LOGISTIK
MULTINOMIAL MENGGUNAKAN METODE *LEAST
ABSOLUTE SHRINKAGE AND SELECTION OPERATOR*
(LASSO)**

**(Studi Kasus Pilihan Sekolah SMA Sederajat Oleh Siswa SMP
dan MTs Negeri di Kabupaten Trenggalek)**

oleh :

**HAFID WAHYU RAMADHAN
135090500111008**

**Setelah dipertahankan di depan Majelis Pengaji
pada tanggal 18 Desember 2017
dan dinyatakan memenuhi syarat untuk memperoleh gelar Sarjana
Sains dalam bidang Statistika**

Dosen Pembimbing

**Achmad Efendi, S.Si, M.Si, Ph.D
NIP. 198102192005011001**

**Mengetahui,
Ketua Jurusan Matematika
Fakultas MIPA Universitas Brawijaya**

**Ratno Bagus Edy Wibowo, S.Si., M.Si., Ph.D.
NIP. 197509082000031003**

LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Hafid Wahyu Ramadhan
NIM : 135090500111008
Jurusan : Matematika
Program Studi : Statistika
Judul Skripsi : Pendugaan Parameter Regresi Logistik Multinomial Menggunakan Metode *Least Absolute Shrinkage And Selection Operator (LASSO)*
(Studi Kasus Pilihan Sekolah Sma Sederajat Oleh Siswa Smp Dan Mts Negeri Di Kabupaten Trenggalek)

Dengan ini menyatakan bahwa:

1. Isi dari Skripsi yang saya buat adalah benar-benar karya sendiri dan tidak menjiplak karya orang lain, selain nama-nama yang termaktub di isi dan tertulis di daftar pustaka dalam Skripsi ini.
2. Apabila dikemudian hari ternyata Skripsi yang saya tulis terbukti hasil jiplakan, maka saya akan bersedia menanggung segala resiko yang akan saya terima.

Demikian pernyataan ini dibuat dengan segala kesadaran.

Malang, 18 Desember 2017
Yang menyatakan,

Hafid Wahyu Ramadhan
135090500111008

**PENDUGAAN PARAMETER REGRESI LOGISTIK
MULTINOMIAL MENGGUNAKAN METODE LEAST
ABSOLUTE SHRINKAGE AND SELECTION OPERATOR
(LASSO)**

**(Studi Kasus Pilihan Sekolah SMA Sederajat Oleh Siswa SMP
dan MTs Negeri di Kabupaten Trenggalek)**

ABSTRAK

Regresi logistik multinomial merupakan analisis yang digunakan untuk mengetahui hubungan antara peubah prediktor dengan peubah respon yang memiliki lebih dari dua kategori. Proses pendugaan parameter dapat dilakukan dengan metode *Maximum Likelihood Estimation* (MLE) dengan pengujian parameter secara parsial dan mereduksi peubah yang tidak berpengaruh. LASSO (*Least Absolute Shrinkage and Selection Operator*) merupakan metode yang dapat menyusutkan koefisien regresi menjadi nol dan sekaligus dapat melakukan pemilihan peubah. LASSO diterapkan dengan menambahkan pinalti LASSO pada fungsi *log likelihood* dalam pendugaan parameter. Data yang digunakan merupakan pilihan sekolah SMA sederajat oleh siswa SMP dan MTs Negeri di Kabupaten Trenggalek. Hasil analisis menggunakan MLE menunjukkan bahwa faktor yang berpengaruh terhadap pilihan adalah nilai rata-rata semester dua saat kelas delapan, jenis sekolah, banyak tanggungan, dan tujuan jangka panjang. Sedangkan faktor yang berpengaruh menggunakan metode LASSO adalah nilai rata-rata semester dua saat kelas delapan, jenis sekolah, pekerjaan ayah, pekerjaan ibu, pendidikan ibu, pendapatan ayah, pendapatan ibu, tujuan jangka panjang dan hal yang dipertimbangkan.

Kata kunci: Regresi logistik multinomial, pemilihan peubah, LASSO, pilihan sekolah

**MULTINOMIAL LOGISTIC REGRESSION PARAMETER
ESTIMATION USING LEAST ABSOLUTE SHRINKAGE
AND SELECTION OPERATOR (LASSO)**

**(Case Study of Senior High School Choice by Public Junior High
School and Islamic Junior High School Students in Trenggalek
Regency)**

ABSTRACT

Multinomial logistic regression is a model used for identifying the relationship between predictor variable and response variable with more than two categories. Parameter estimation process can be done using Maximum Likelihood Estimation (MLE) with partial significance test followed by reducing non influencing variables. LASSO (Least Absolute Shrinkage and Selection Operator) is a method that capable of shrinking regression coefficient into zero and selecting variables. LASSO is applied by adding LASSO penalty into log likelihood function in parameter estimation. The data are senior high school choices by Public Junior High School And Islamic Junior High School students in Trenggalek Regency. The result showed that influencing factors are average score during eight semester, school type, dependents of family, and long term plan. Meanwhile the influencing factors from model with LASSO method are average score during eight semester, school type, father's occupation, mother's occupation, mother's last education, father's income, mother's income, long term plan and the most considered aspect.

Keywords: Multinomial logistic regression, variable selection, LASSO, school choice

KATA PENGANTAR

Puji syukur atas kehadiran Allah SWT yang telah memberikan rahmat serta hidayahnya sehingga penulis dapat menyelesaikan skripsi dengan judul “Pendugaan Parameter Regresi Logistik Multinomial dengan Metode *Least Absolute Shrinkage and Selection Operator (LASSO)*” sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains dalam bidang statistika. Dalam penulisan skripsi ini, penulis telah dibantu oleh berbagai pihak. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih kepada :

1. Bapak Achmad Efendi, S.Si, M.Si, Ph.D selaku Dosen Pembimbing Skripsi yang telah berkenan meluangkan waktu, ilmu dan tenaga dalam memberikan arahan selama penyusunan penulisan skripsi.
2. Ibu Rahma Fitriani, S.Si, M.Si, Ph.D selaku Dosen Pengaji 1 dan Ibu Nurjannah, S.Si, M.Sc, Ph.D selaku Dosen Pengaji 2 atas segala nasihat yang telah diberikan selama penyusunan skripsi.
3. Bapak Ratno Bagus Edy Wibowo, S.Si, M.Si, Ph.D selaku Ketua Jurusan Matematika FMIPA Universitas Brawijaya Malang.
4. Semua karyawan Jurusan Matematika FMIPA Universitas brawijaya.
5. Ibu Munis dan Bapak Redi Susiloadi atas kasih sayang, optimisme serta doa yang selalu diberikan.
6. Teman-teman seperjuangan Statistika angkatan 2013 atas kerjasama dan kebersamaannya selama ini.
7. Destian, Yuyud, Rauzan, dan seluruh anggota skuat *Lovable Project* yang banyak membantu saya belajar selama penggerjaan skripsi.
8. Keluarga besar HIMAMASTA, khususnya kabinet Masta Sapa, Masta Slam dan Masta Secant.
9. Seluruh siswa, kepala, guru karyawan SMP dan MTs Negeri di Kabupaten Trenggalek.
10. Semua pihak yang telah banyak membantu dan memberikan semangat selama penulisan skripsi ini.

Penulis menyadari bahwa dalam penulisan skripsi ini masih terdapat kekurangan. Oleh karena itu, penulis mengharapkan kritik maupun saran dari pembaca yang dapat membangun untuk penulisan yang lebih baik. Semoga skripsi ini bermanfaat untuk berbagai pihak yang membutuhkan.

Malang, Desember 2017

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN SKRIPSI	ii
LEMBAR PERNYATAAN	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	xi
BAB I PENDAHULUAN	
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan Penelitian.....	2
1.4. Manfaat Penelitian.....	3
BAB II TINJAUAN PUSTAKA	
2.1. Regresi Logistik	5
2.2. Regresi Logistik Multinomial	5
2.3. Multikolinieritas	6
2.4. Pendugaan Regresi Logistik Multinomial	7
2.5. Pengujian Parameter	9
2.5.1. Pengujian parameter secara simultan	10
2.5.2. Pengujian parameter secara parsial	10
2.6. Uji Kesesuaian Model	11
2.7. <i>Least Absolute Shrinkage and Selection Operator</i>	12
2.8. Interpretasi Parameter	14
2.9. Pengambilan Sampel	15
2.10. Pendidikan	16
2.11. Sekolah Lanjutan SMA, SMK dan MA	16
2.12. Faktor-faktor yang mempengaruhi pilihan sekolah	17
BAB III METODE PENELITIAN	
3.1. Sumber Data	21
3.2. Metode Penarikan Sampel	22
3.3. Metode Analisis.....	25
BAB IV HASIL DAN PEMBAHASAN	
4.1. Analisis Deskriptif	29
4.2. Pendekripsi Multikolinieritas	30
4.3. Regresi Logistik Multinomial	30
4.3.1. Pendugaan Parameter	30
4.3.2. Pengujian Parameter Secara Simultan	32

4.3.3. Pengujian Parameter Secara Parsial	33
4.3.4. Pengujian Kesesuaian Model	35
4.4. Regresi Logistik Multinomial dengan LASSO	35
4.5. Interpretasi Model	39
4.5.1. Interpretasi Model Regresi Logistik Multinomial	39
4.5.2. Interpretasi Model Regresi Logistik Multinomial dengan Metode LASSO	44
BAB V KESIMPULAN DAN SARAN	
5.1. Kesimpulan	49
5.2. Saran	49
DAFTAR PUSTAKA	51
LAMPIRAN	53

DAFTAR TABEL

3.1. Peubah Prediktor	21
3.2. Sampel Kelas	23
3.3. Banyak Siswa tiap Kelas.....	23
4.1. Nilai VIF untuk masing-masing prediktor	30
4.2. Penduga Parameter	31
4.3. Hasil Pengujian Parameter secara Simultan	32
4.4. Hasil Pengujian Parameter secara Parsial	32
4.5. Hasil pengujian Kesesuaian Model	35
4.6. Penduga Parameter menggunakan LASSO	35
4.7. Parameter Berpengaruh Menggunakan LASSO	36
4.8. Perbandingan Perbandingan Koefisien yang berpengaruh antara metode MLE dan Metode Lasso	37
4.9. Nilai AIC	38
4.10. <i>Odds Ratio</i> Logit ke-1 (SMA) dengan Metode <i>Maximum Likelihood</i>	40
4.11. <i>Odds Ratio</i> Logit ke-2 (SMK) dengan Metode <i>Maximum Likelihood</i>	40
4.12. <i>Odds Ratio</i> Logit ke-3 (MA) dengan Metode <i>Maximum Likelihood</i>	41
4.14. Perbandingan <i>Odds Ratio</i> untuk Prediktor yang Sama (MLE).....	42
4.14. <i>Odds Ratio</i> Logit ke-1 (SMA) dengan Metode LASSO	44
4.15. <i>Odds Ratio</i> Logit ke-2 (SMK) dengan Metode LASSO	46
4.16. <i>Odds Ratio</i> Logit ke-2 (MA) dengan Metode LASSO	47
4.17. Perbandingan <i>Odds Ratio</i> untuk Prediktor yang Sama (LASSO)	47

DAFTAR GAMBAR

3.1. Diagram Alir Penelitian	26
4.1. Pilihan Sekolah Lanjutan Siswa SMP dan MTs Negeri di Kabupaten Trenggalek	29

DAFTAR LAMPIRAN

Lampiran 1.	Angket	53
Lampiran 2.	Tabulasi Hasil Penyebaran Angket Pilihan Sekolah Siswa SMP dan MTS Negeri di Kabupaten Trenggalek	55
Lampiran 3.	Data Pilihan Sekolah Siswa SMP dan MTs Negeri di Kabupaten Trenggalek.....	59
Lampiran 4.	Grafik Perbandingan Pilihan Perdasarkan Masing-Masing Peubah.....	61
Lampiran 5.	Surat Keterangan Pengambilan Data di Sekolah	67
Lampiran 6.	Nilai VIF	79
Lampiran 7.	Output SPSS 22 untuk Persentase	83
Lampiran 8.	Output SPSS 22 untuk Uji Simultan dan Uji Kesesuaian Model.....	85
Lampiran 9.	Output SPSS 22 untuk Pendugaan Parameter, Uji Parsial dan <i>Odds Ratio</i>	87
Lampiran 10.	Syntax R untuk Pendugaan Regresi Logistik Multinomial menggunakan LASSO	97
Lampiran 11.	Output R Pendugaan Parameter Regresi Logistik Multinomial menggunakan LASSO	99
Lampiran 12.	Perbandingan Penduga Parameter Regresi Logistik Multinomial menggunakan LASSO	103

BAB I

PENDAHULUAN

1.1. Latar Belakang

Pendidikan merupakan salah satu aspek penting dalam pembangunan di Indonesia. Sumber daya manusia yang berkualitas didukung oleh pendidikan yang baik. Salah satu usaha pemerintah dalam meningkatkan kualitas sumberdaya manusia adalah mencanangkan wajib belajar (WAJAR) 9 tahun hingga SMP sederajat. Beberapa daerah seperti Kabupaten Trenggalek sudah merencanakan untuk mencanangkan WAJAR 12 tahun. Kebijakan tersebut mengisyaratkan bahwa setiap warga negara berpendidikan minimal jenjang SMP atau sederajat. Bagi Kabupaten Trenggalek yang membuat target WAJAR 12 tahun, Sekolah Menengah Atas (SMA) atau sederajat menjadi prioritas pendidikan yang penting untuk dicapai. Berdasarkan data proyeksi pendidikan oleh Kementerian Kebudayaan dan Pendidikan Dasar (Kemendikbud) pada tahun 2013 menunjukkan bahwa dari tahun 2007 sampai 2011 terjadi peningkatan lulusan SMP dan MTs yang tidak melanjutkan pendidikan sekitar 12,20 % per tahun. Pada 2012 sampai tahun 2020, Kemendikbud memproyeksikan bahwa angka tersebut akan berubah sehingga lulusan SMP dan MTs yang tidak lanjut akan menurun hingga 16,66 % per tahun.

Banyaknya lembaga pendidikan SMA sederajat yang adaseperti Sekolah Menengah Atas (SMA), Sekolah Menengah Kejuruan (SMA) dan Madrasah Aliyah (MA) menjadi permasalahan untuk siswa-siswi SMP dan MTs kelas VIII dalam memilih pendidikan berikutnya. Banyak terdapat siswa yang memilih pendidikan tingkat SMA sederajat tanpa memperhatikan minat dan kemampuan diri sendiri. Banyak faktor yang mempengaruhi pilihan siswa dalam menentukan sekolah yang akan dimasuki. Dalam penelitian ini, faktor yang digunakan secara khusus adalah faktor akademik dan ekonomi diantaranya nilai, jumlah saudara, serta pendidikan dan pekerjaan orangtua. Faktor-faktor tersebut perlu diketahui oleh siswa, orang tua, dan pihak sekolah untuk memaksimalkan minat dan kemampuan saat memilih pendidikan setelah SMP. Analisis statistika dilakukan untuk mengetahui faktor yang berpengaruh terhadap pilihan siswa-siswi SMP dan MTs untuk memilih jenjang sekolah yang lebih tinggi. Analisis statistika yang tepat digunakan untuk bidang ini adalah analisis regresi logistik.

Regresi logistik digunakan untuk mengetahui hubungan antara peubah respon berupa data kategorik berskala nominal maupun ordinal dengan satu atau lebih peubah prediktor kontinyu atau kategorik. Untuk membentuk model regresi diperlukan pendugaan parameter. Regresi logistik yang digunakan untuk peubah respon dengan lebih dari dua

kategori disebut regresi logistik multinomial. Dalam kasus pilihansekolah, terdapat banyak faktor yang berpotensi mempengaruhi pilihan siswa. Oleh karena itu dibutuhkan metode yang dengan cepat menyeleksi faktor yang berpengaruh. Dalam penelitian ini digunakan pendugaan parameter *Least Absolute Shrinkage and Selection Operator* (LASSO). Metode ini dapat menyusutkan koefisien regresi menjadi nol serta melakukan pemilihan peubah prediktor. Metode ini juga dapat digunakan saat terjadi kasus multikolinieritas di mana terdapat korelasi yang tinggi antar peubah prediktor. Dalam penelitian Aziz (2016), metode ini telah digunakan pada regresi logistik ordinal. Pendugaan menggunakan metode LASSO digunakan dalam penelitian ini sehingga bisa langsung menentukan faktor yang berpengaruh dari banyak faktor yang ada tanpa melalui proses panjang pengujian parameter parsial maupun simultan. Dalam penelitian ini, variabel faktor-faktor pilihan siswa SMP dan MTs di Negeri di Kabupaten Trenggalek dalam melanjutkan pendidikan sederajat SMA berupa peubah bersifat kontinu dan kategorik.

1.2. Rumusan Masalah

1. Bagaimana model regresi logistik multinomial dengan metode LASSO untuk pilihan sekolah SMA sederajat siswa SMP dan MTs Negeri di Kabupaten Trenggalek?
2. Apa saja faktor-faktor yang berpengaruh secara signifikan untuk pilihan SMA sederajat siswa SMP dan MTs Negeri di Kabupaten Trenggalek menggunakan regresi logistik multinomial dengan metode LASSO?

1.3. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan tersebut, maka tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut

1. Membentuk model regresi logistik multinomial dengan metode LASSO untuk pilihan SMA sederajat siswa SMP dan MTs Negeri di Kabupaten Trenggalek.
2. Menentukan faktor-faktor yang berpengaruh secara signifikan terhadap pilihan SMA sederajat siswa SMP dan MTs Negeri di Kabupaten Trenggalek menggunakan regresi logistik multinomial dengan metode LASSO?

1.4. Manfaat Penelitian

Manfaat yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Memberikan tambahan informasi mengenai analisis regresi logistik multinomial dengan menggunakan pendugaan parameter dengan

metode LASSO dan mendapatkan model dari pendugaan tersebut sehingga dapat digunakan untuk mengetahui faktor-faktor yang mempengaruhi pilihan SMA sederajat siswa SMP dan MTs Negeri di Kabupaten Trenggalek.

2. Menambah pengetahuan para siswa, orangtua serta pihak SMP dan MTs Negeri di Kecamatan Munjungan Kabupaten Trenggalek mengenai faktor-faktor yang berpengaruh dalam pemilihan sekolah setingkat SMA.

1.5. Batasan Masalah

Batasan masalah yang digunakan pada penelitian ini adalah sebagaimana berikut:

1. Penelitian ini membahas tentang penerapan regresi logistik multinomial menggunakan pendugaan parameter dengan metode LASSO pada kasus pilihan SMA sederajat siswa SMP dan MTs Negeri di Kabupaten Trenggalek.
2. Data yang diambil merupakan data primer dengan peubah prediktor yang digunakan yaitu nilai rata-rata kelas VIII, pekerjaan ayah, pekerjaan ibu, pendidikan ayah, pendidikan ibu, pendapatan per bulan ayah, pendapatan per bulan ibu, banyak tanggungan, tujuan jangka panjang, hal penting yang dipertimbangkan, dan pihak yang paling mempengaruhi pilihan.

BAB II

TINJAUAN PUSTAKA

2.1. Regresi Logistik

Pada umumnya, pada analisis regresi peubah respon yang digunakan memiliki skala interval dan rasio, tetapi pada kasus-kasus terapan seperti penelitian ini digunakan peubah respon dengan skala kategorik yaitu skala nominal maupun ordinal. Untuk melakukan analisis pada data kategorik, analisis regresi yang sesuai untuk digunakan adalah analisis regresi logistik.

Menurut Hosmer dan Lemeshow (2000), regresi logistik dapat digunakan untuk menggambarkan hubungan peubah prediktor dan peubah respon yang memiliki dua atau lebih kategori dengan peubah prediktor yang dapat bersifat kontinyu, kategorik maupun campuran keduanya.

Pada umumnya, model regresi logistik memiliki peubah respon dengan dua kategori atau biasa disebut regresi logistik biner. Pada penerapan sehari-hari, terdapat banyak kasus di mana dibutuhkan respon yang memiliki lebih dari dua kategori. Model regresi seperti ini disebut sebagai model regresi logistik *polytomous*. Model regresi ini dapat digunakan pada peubah respon dengan skala nominal atau skala ordinal.

2.2. Regresi Logistik Multinomial

Regresi logistik multinomial merupakan pengembangan dari regresi logistik untuk peubah respon yang memiliki lebih dari dua kategori. Model regresi ini dapat digunakan untuk menyelesaikan kasus yang memiliki peubah respon berskala nominal dengan level lebih dari dua kategori dengan peubah prediktor yang bersifat kategori atau kontinu.

Menurut Hosmer dan Lemeshow (2000) bentuk persamaan regresi logistik multinomial adalah sebagai berikut:

$$\pi_j(x_i) = \frac{e^{g_j(x_i)}}{1+e^{g_1(x_i)}+e^{g_2(x_i)}+\dots+e^{g_j(x_i)}} \quad (2.1)$$

Dari persamaan (2.1) dapat dituliskan dalam *log odds* atau logit. Logit merupakan log natural (*ln*) dari *odds*. Dimana *odds* merupakan peluang kejadian sukses dibandingkan dengan peluang gagal. Secara umum fungsi logit peubah respon ke-*j* pada peubah prediktor ke-*i* dinyatakan sebagai berikut:

$$g_j(x_i) = \ln \left(\frac{\pi_j(x_i)}{1-\pi_j(x_i)} \right) = \beta_{j0} + \beta_{j1}x_1 + \beta_{j2}x_2 + \dots + \beta_{jp}x_{pi}$$
$$g_j(x_i) = \beta_{j0} + \sum_{k=1}^p \beta_{jk} x_{ki} \quad (2.2)$$

di mana

x_i : vektor yang memuat nilai-nilai pengamatan ke-*i* dari peubah prediktor ke-*j*

$\pi_j(x_i)$: peluang kategori respon ke-*j* pada peubah prediktor ke-*i*

- $g_j(x_i)$: fungsi logit peubah respon untuk kategori ke- j pada peubah prediktor ke- i
 β_{jk} : koefisien model kategori ke- j peubah prediktor ke- k
 x_{ki} : nilai peubah prediktor ke- k pengamatan ke- i
 j : indeks untuk kategori ($j = 0, 1, 2, \dots, j-1$)
 k : indeks untuk peubah prediktor ($k = 0, 1, 2, \dots, p$)
 i : indeks untuk pengamatan ($i = 0, 1, 2, \dots, n$)

Apabila terdapat j kategori pada peubah respon, maka akan terdapat $j-1$ fungsi logit. Hal ini disebabkan oleh salah satu kategori yang dijadikan pembanding yaitu kategori pertama. Dengan menggunakan transformasi logit sesuai persamaan (2.3) maka didapatkan tiga fungsi logit dari peubah respon dengan empat kategori. Fungsi logit pertama adalah sebagai berikut

$$g_1(x_i) = \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \dots + \beta_{1k}x_{ki} \quad (2.3)$$

Dari fungsi logit pada persamaan (2.1), apabila $\pi_j(x_i) = P(Y = j|x_i)$; ($j = 0, 1, 2, 3$) menyatakan peluang kategori respon ke- j pada peubah prediktor ke- i dengan $e^{g_0(x_i)} = 1$ maka, apabila terdapat empat kategori model regresi logistik multinomial pada peubah respon ke- j adalah:

Peluang kategori peubah respon pertama

$$\pi_0(x_i) = \frac{e^{g_0(x_i)}}{1+e^{g_1(x_i)}+e^{g_2(x_i)}+e^{g_3(x_i)}} = \frac{1}{1+e^{g_1(x_i)}+e^{g_2(x_i)}+e^{g_3(x_i)}} \quad (2.4)$$

Peluang kategori peubah respon kedua

$$\pi_1(x_i) = \frac{e^{g_1(x_i)}}{1+e^{g_1(x_i)}+e^{g_2(x_i)}+e^{g_3(x_i)}} \quad (2.5)$$

Peluang kategori peubah respon ketiga

$$\pi_2(x_i) = \frac{e^{g_2(x_i)}}{1+e^{g_1(x_i)}+e^{g_2(x_i)}+e^{g_3(x_i)}} \quad (2.6)$$

Peluang kategori peubah respon keempat

$$\pi_3(x_i) = \frac{e^{g_3(x_i)}}{1+e^{g_1(x_i)}+e^{g_2(x_i)}+e^{g_3(x_i)}} \quad (2.7)$$

2.3. Multikolinieritas

Multikolinieritas merupakan hubungan linier antara beberapa peubah atau semua peubah prediktor pada model regresi. Pada regresi logistik yang menggunakan lebih dari satu peubah prediktor perlu dilakukan pengujian terhadap multikolinieritas. Pembentukan model regresi logistik juga sensitif terhadap multikolinieritas. Multikolinieritas menyebabkan besarnya nilai *standard error* (SE) dari pendugaan koefisien regresi (Gujarati, 2004). Apabila nilai SE semakin besar, nilai statistik uji *Wald* pengujian koefisien regresi logistik menjadi semakin kecil. Hal tersebut menyebabkan semakin kecilnya kemungkinan peubah prediktor berpengaruh terhadap peubah respon. Beberapa cara untuk mendeteksi multikolinieritas antara lain

tingginya korelasi antar peubah prediktor, tingginya nilai R^2 namun peubah prediktor yang berpengaruh terhadap peubah respon sedikit, adanya korelasi yang tinggi antar peubah prediktor dan dengan melihat nilai *Variance Inflation Factor* (VIF).

$$(VIF)_k = \frac{1}{1 - R_k^2} \quad (2.8)$$

$$k = 1, 2, \dots, p$$

di mana:

p : banyaknya peubah prediktor

R_k^2 : koefisien determinasi ganda jika X_k diregresikan dengan $p - 1$ peubah prediktor lain

Koefisien determinasi diperoleh dari model *auxiliary regression* yaitu regresi masing-masing peubah prediktor dengan seluruh peubah prediktor. Salah satu peubah prediktor berperan sebagai peubah respon dan peubah prediktor lainnya akan berperan sebagai peubah prediktor. Apabila peubah prediktor yang berperan sebagai peubah respon berskala interval atau rasio maka koefisien determinasi yang digunakan merupakan koefisien determinasi model regresi linier berganda.

$$R^2 = \frac{JKR}{JKT} = 1 - \frac{JKG}{JKT} \quad (2.9)$$

Sedangkan apabila peubah prediktor yang berperan sebagai peubah respon berskala nominal atau ordinal, maka koefisien determinasi yang digunakan adalah koefisien determinasi model regresi logistik, yaitu $R_{Cox-Snell}^2$. Koefisien determinasi Cox-Snell dapat dituliskan sebagai:

$$R_{Cox-Snell}^2 = 1 - \left(\frac{L_0}{L_1} \right)^{2/n} \quad (2.10)$$

di mana:

L_0 : nilai *likelihood* model tereduksi

L_1 : nilai *likelihood* model penuh

n : banyaknya pengamatan

Nilai VIF semakin meningkat apabila terdapat korelasi yang besar antar peubah prediktor. Asumsi non multikolinieritas tidak terpenuhi apabila nilai $VIF_k > 10$.

2.4. Pendugaan Regresi Logistik Multinomial

Metode yang sering digunakan dalam pendugaan parameter regresi logistik adalah metode *Maximum Likelihood*. Menurut Kleinbaum dan Klein (2010) alasan digunakannya metode *Maximum Likelihood* karena metode tersebut dapat digunakan untuk pendugaan model linier ataupun non linier. Pada regresi logistik multinomial peubah respon bersifat multinomial sehingga fungsi *likelihood* sesuai dengan sebaran multinomial.

Menurut Hosmer dan Lemeshow (2000) parameter yang dipilih adalah parameter yang memaksimumkan fungsi kemungkinan (*likelihood*),

sehingga parameter yang diperoleh merupakan parameter yang mendekati data yang diamati. Bentuk umum fungsi pendugaan parameter menggunakan metode *Maximum Likelihood* adalah:

$$L(\beta) = \prod_{i=1}^n [\pi_0(x_i)^{y_{0i}} \pi_1(x_i)^{y_{1i}} \dots \pi_{j-1}(x_i)^{y_{(j-1)i}}], \quad (2.11)$$

di mana $\beta = (\beta_{j0}, \beta_{j1}, \dots, \beta_{jp})$

Sehingga logaritma natural dari fungsi *likelihood* pada persamaan 2.11 adalah

$$l(\beta) = \ln \prod_{i=1}^n [\pi_0(x_i)^{y_{0i}} \pi_1(x_i)^{y_{1i}} \dots \pi_{j-1}(x_i)^{y_{(j-1)i}}], \quad (2.12)$$

di mana:

y_{ij} =banyaknya peubah respon kategori ke- j dan pengamatan ke- i .

Menurut Hosmer (2000), dalam membuat fungsi *likelihood* pada model dengan peubah respon lebih dari dua kategori, terlebih dahulu membuat peubah respon biner. Peubah ini digunakan untuk menunjukkan anggota dari kelompok pada pengamatan. Misal terdapat 4 kategori peubah respon, maka akan terbentuk peubah biner dengan $y_{ji}=(y_{0i}, y_{1i}, y_{2i}, \dots, y_{ji})$, yaitu:

$y_{ji}=0$ maka $y_{0i}=1, y_{1i}=0, y_{2i}=0, y_{3i}=0$

$y_{ji}=1$ maka $y_{0i}=0, y_{1i}=1, y_{2i}=0, y_{3i}=0$

$y_{ji}=2$ maka $y_{0i}=0, y_{1i}=0, y_{2i}=1, y_{3i}=0$

$y_{ji}=3$ maka $y_{0i}=0, y_{1i}=0, y_{2i}=0, y_{3i}=1$

Menggunakan persamaan 2.10 dan 2.11, maka didapatkan fungsi *likelihood* untuk $j=4$

$$L(\beta) = \prod_{i=1}^n [\pi_0(x_i)^{y_{0i}} \pi_1(x_i)^{y_{1i}} \pi_2(x_i)^{y_{2i}} \pi_3(x_i)^{y_{3i}}] \quad (2.13)$$

Logaritma fungsi likelihood dari persamaan 2.13 adalah:

$$\begin{aligned} l(\beta) &= \ln \prod_{i=1}^n [\pi_0(x_i)^{y_{0i}} \pi_1(x_i)^{y_{1i}} \pi_2(x_i)^{y_{2i}} \pi_3(x_i)^{y_{3i}}] \\ &= \ln \prod_{i=1}^n \left[\left(\frac{1}{1 + e^{g_1(x_i)} + e^{g_2(x_i)} + e^{g_3(x_i)}} \right)^{y_{0i}} \left(\frac{e^{g_1(x_i)}}{1 + e^{g_1(x_i)} + e^{g_2(x_i)} + e^{g_3(x_i)}} \right)^{y_{1i}} \dots \left(\frac{e^{g_3(x_i)}}{1 + e^{g_1(x_i)} + e^{g_2(x_i)} + e^{g_3(x_i)}} \right)^{y_{3i}} \right] \\ &= \ln \prod_{i=1}^n \left[(e^{g_1(x_i)})^{y_{1i}} (e^{g_2(x_i)})^{y_{2i}} (e^{g_3(x_i)})^{y_{3i}} \left(\frac{1}{1 + e^{g_1(x_i)} + e^{g_2(x_i)} + e^{g_3(x_i)}} \right)^{y_{0i}+y_{1i}+y_{2i}+y_{3i}} \right] \end{aligned} \quad (2.14)$$

Jika diketahui bahwa $\sum_{i=1}^n y_{ji} = 1$, maka didapatkan logaritma natural dari fungsi *likelihood*:

$$\begin{aligned} l(\beta) &= \ln \prod_{i=1}^n \left[(e^{g_1(x_i)})^{y_{1i}} (e^{g_2(x_i)})^{y_{2i}} (e^{g_3(x_i)})^{y_{3i}} \left(\frac{1}{1 + e^{g_1(x_i)} + e^{g_2(x_i)} + e^{g_3(x_i)}} \right)^{y_{0i}+y_{1i}+y_{2i}+y_{3i}} \right] \\ &= \ln \left[(e^{g_1(x_i)})^{y_{1i}} (e^{g_2(x_i)})^{y_{2i}} (e^{g_3(x_i)})^{y_{3i}} (1 + e^{g_1(x_i)} + e^{g_2(x_i)} + e^{g_3(x_i)})^{-\sum_{i=1}^n (y_{0i}+y_{1i}+y_{2i}+y_{3i})} \right] \\ &= \ln(e^{g_1(x_i)})^{y_{1i}} + \ln(e^{g_2(x_i)})^{y_{2i}} + \ln(e^{g_3(x_i)})^{y_{3i}} + \ln(1 + e^{g_1(x_i)} + e^{g_2(x_i)} + e^{g_3(x_i)})^{-\sum_{i=1}^n (y_{0i}+y_{1i}+y_{2i}+y_{3i})} \\ &= \sum_{i=1}^n y_{1i} \ln(e^{g_1(x_i)}) + \sum_{i=1}^n y_{2i} \ln(e^{g_2(x_i)}) + \sum_{i=1}^n y_{3i} \ln(e^{g_3(x_i)}) - \sum_{i=1}^n (y_{0i} + y_{1i} + y_{2i} + y_{3i}) \ln(1 + e^{g_1(x_i)} + e^{g_2(x_i)} + e^{g_3(x_i)}) \end{aligned} \quad (2.15)$$

$l(\beta)$ dapat dituliskan seperti pada persamaan 2.16 berikut:

$$l(\beta) = \sum_{i=1}^n y_{1i} \ln(e^{\theta_1(x_i)}) + \sum_{i=1}^n y_{2i} \ln(e^{\theta_2(x_i)}) + \sum_{i=1}^n y_{3i} \ln(e^{\theta_3(x_i)}) - \ln(1 + e^{\theta_1(x_i)} + e^{\theta_2(x_i)} + e^{\theta_3(x_i)})$$

$$l(\beta) = \sum_{i=1}^n \left[[y_{1i}g_1(x_i) + y_{2i}g_2(x_i) + y_{3i}g_3(x_i)] - \ln(1 + e^{\theta_1(x_i)} + e^{\theta_2(x_i)} + e^{\theta_3(x_i)}) \right] \quad (2.16)$$

Untuk memperoleh nilai β_{jk} yang memaksimumkan $l(\beta)$, maka dilakukan turunan pertama terhadap β_{jk} .

$$\begin{aligned} \frac{\partial l(\beta)}{\partial \beta_{jk}} &= \sum_{i=1}^n (y_{1i}(x_{0i} + x_{1i} + x_{2i} + \dots + x_{pi}) + y_{2i}(x_{0i} + x_{1i} + x_{2i} + \dots + x_{pi}) \\ &\quad + y_{3i}(x_{0i} + x_{1i} + x_{2i} + \dots + x_{pi}) \\ &\quad - \frac{1}{1 + e^{g_1(x_i)} + e^{g_2(x_i)} + e^{g_3(x_i)}} [e^{g_1(x_i)(x_{0i} + x_{1i} + x_{2i} + \dots + x_{pi})} \\ &\quad + e^{g_2(x_i)(x_{0i} + x_{1i} + x_{2i} + \dots + x_{pi})} + e^{g_3(x_i)(x_{0i} + x_{1i} + x_{2i} + \dots + x_{pi})}] \\ &= \sum_{i=1}^n x_{ki} y_{ji} - \frac{1}{1 + e^{g_1(x_i)} + e^{g_2(x_i)} + e^{g_3(x_i)}} [e^{g_1(x_i)(x_{0i} + x_{1i} + x_{2i} + \dots + x_{pi})} \\ &\quad + e^{g_2(x_i)(x_{0i} + x_{1i} + x_{2i} + \dots + x_{pi})} + e^{g_3(x_i)(x_{0i} + x_{1i} + x_{2i} + \dots + x_{pi})}] \\ \frac{\partial l(\beta)}{\partial \beta_{jk}} &= \sum_{i=1}^n x_{ki} y_{ji} - \sum_{i=1}^n x_{ki} \pi_{ji} \\ \frac{\partial l(\beta)}{\partial \beta_{jk}} &= \sum_{i=1}^n x_{ki} (y_{ji} - \pi_{ji}) \end{aligned} \tag{2.17}$$

di mana:

x_{ki} : nilai peubah prediktor ke- k pada pengamatan ke- i

y_{ji} : nilai peubah respon kategori ke- j pada pengamatan ke- i

π_{ji} : nilai peluang kategori respon ke- j pada pengamatan ke- i

Hasil turunan parsial pertama fungsi log *likelihood* bersifat nonlinier maka untuk menyelesaikan persamaan 2.17 dengan solusi penduga parameter yang memaksimumkan *likelihood* diperlukan iterasi dari metode numerik. Metode Numerik yang dapat digunakan adalah Newton-Raphson. Berikut adalah tahapan iterasi Newton-Raphson:

1. Menentukan Penduga awal $\beta_{jk}^t, \beta_{jk}^1 = 0$, t merupakan tahapan iterasi.
2. Melakukan iterasi untuk mendapatkan penduga β_{jk}^t yang baru:

$$\beta_{jk}^{t+1} = \beta_{jk}^t - (\mathbf{H}(\beta_{jk}^t))^{-1} l'(\beta_{jk}^t)$$
di mana :
3. Iterasi selesai apabila $\beta_{jk}^{t+1} \approx \beta_{jk}^t$ atau $\beta_{jk}^{t+1} - \beta_{jk}^t \approx 0$

2.5. Pengujian Parameter

Pengujian parameter dilakukan secara simultan menggunakan *likelihood ratio test* (Uji G) dan secara parsial menggunakan statistik uji *Wald*.

2.5.1. Pengujian parameter secara simultan

Pengujian secara simultan dengan *likelihood ratio test* (Uji G) digunakan untuk memeriksa apakah peubah prediktor berpengaruh terhadap peubah respon. Hipotesis yang melandasi pengujian ini adalah

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0 ; \beta_{jk} = 0$$

vs

$$H_1 : \text{paling tidak terdapat satu } k \text{ di mana } \beta_{jk} \neq 0$$

Jika H_0 benar maka statistik uji χ^2 likelihood ratio:

$$G = -2 \log \left(\frac{L_0}{L_1} \right) = -2(\log(L_0) - \log(L_1)) \sim \chi^2_p \quad (2.18)$$

di mana:

p : banyaknya peubah prediktor dalam model

L_0 : nilai likelihood model tereduksi

L_1 : nilai likelihood model penuh (Kleinbaum dan Klein. 2010)

Statistik uji likelihood ratio juga dapat dituliskan sebagai

$$G = -2 \log \left[\frac{\left(\frac{n_{1i}}{n} \right)^{n_{1i}} \left(\frac{n_{2i}}{n} \right)^{n_{2i}} \dots \left(\frac{n_{ji}}{n} \right)^{n_{ji}}}{\prod_{i=1}^n \pi_1(x_i)^{y_{1i}} \pi_2(x_i)^{y_{2i}} \dots \pi_j(x_i)^{y_{ji}}} \right] \sim \chi^2_p \quad (2.19)$$

di mana:

n_j : banyaknya individu pada kategori ke- j

$\pi_j(x_i)$: peluang Y kategori ke- j pengamatan ke- i

y_{ji} : nilai peubah respon pada kategori ke- j pengamatan ke- i

H_0 ditolak apabila $G > \chi^2_{(p,a)}$ atau apabila $P(\chi^2_p > G) < \alpha$, sehingga dapat disimpulkan paling sedikit terdapat satu β_k variabel prediktor yang berpengaruh terhadap variabel respon (Hosmer dan Lemeshow, 2000).

2.5.2. Pengujian parameter secara parsial

Pengujian secara parsial dengan uji *Wald* bertujuan untuk memeriksa pengaruh masing-masing koefisien regresi setiap peubah prediktor secara parsial yang terdapat dalam model. Hipotesis yang melandasi uji *Wald* adalah

$$H_0 : \beta_k = 0 \quad ; \text{ peubah prediktor ke-}k \text{ tidak berpengaruh terhadap peubah respon}$$

vs

$$H_1 : \beta_k \neq 0 \quad ; \text{ peubah prediktor ke-}k \text{ berpengaruh terhadap peubah respon}$$

Jika H_0 benar statistik uji:

$$W = \frac{\hat{\beta}_k - \beta_k}{SE(\hat{\beta}_k)} \sim N(0,1) \quad (2.20)$$

$$k = 1, 2, \dots, p$$

di mana :

p : banyaknya peubah prediktor dalam model

$\hat{\beta}_k$: penduga parameter β_k

$SE(\hat{\beta}_k)$: salah baku penduga parameter β_k

$$SE(\hat{\beta}_k) = \sqrt{var(\hat{\beta}_k)} \quad (2.21)$$

$$var(\hat{\beta}_k) = diag[\mathbf{X}'\mathbf{V}\mathbf{X}]^{-1} \quad (2.22)$$

di mana \mathbf{X} merupakan matriks berukuran $n \times (p+1)$ dan \mathbf{V} merupakan matriks diagonal $\hat{\pi}_i(1 - \hat{\pi}_i)$ yang berukuran $n \times n$.

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{bmatrix}$$
$$\mathbf{V} = \begin{bmatrix} \hat{\pi}_1(1 - \hat{\pi}_1) & 0 & \cdots & 0 \\ 0 & \hat{\pi}_2(1 - \hat{\pi}_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{\pi}_n(1 - \hat{\pi}_n) \end{bmatrix}$$

H_0 ditolak apabila $|W| > Z_{\alpha/2}$ atau $P(|Z| > W) < \alpha$, sehingga dapat disimpulkan ada pengaruh antara peubah prediktor dengan peubah respon (Hosmer dan Lemeshow, 2000). Menurut Agresti (2007) apabila statistik uji *Wald* dikuadratkan (W^2) akan mengikuti sebaran khi-kuadrat (χ^2) dengan derajat bebas satu.

2.6. Uji Kesesuaian Model

Uji Kesesuaian model digunakan untuk mengetahui apakah model yang didapatkan dari pendugaan parameter dapat menggambarkan data. Uji yang digunakan adalah statistik uji *Pearson*, dengan hipotesis:

H_0 : Model sesuai

vs

H_1 : Model tidak sesuai

Selisih antara nilai pengamatan dengan penduga (*Pearson residual*) didefinisikan sebagai

$$r(y_j, \hat{\pi}_j) = \frac{(y_j - n_j \pi_j)}{\sqrt{n_j \pi_j (1 - \pi_j)}} \quad (2.23)$$

Sehingga berdasarkan persamaan (2.16) dapat diperoleh statistik uji *Pearson chi-square* jika H_0 benar sebagai berikut

$$\begin{aligned} \chi^2_{Pearson} &= \sum_{j=1}^J r(y_j, \pi_j)^2 \\ &= \sum_{j=1}^J \left(\frac{(y_j - n_j \pi_j)}{\sqrt{n_j \pi_j (1 - \pi_j)}} \right)^2 \sim \chi^2_{n-p-1} \\ \chi^2_{Pearson} &= \sum_{j=1}^J \frac{(y_j - n_j \pi_j)^2}{n_j \pi_j (1 - \pi_j)} \sim \chi^2_{n-p-1} \end{aligned} \quad (2.24)$$

di mana:

y_j : nilai peubah respon pada kategori ke- j

n_j : banyaknya pengamatan pada kategori ke- j

π_j : peluang Y kategori ke- j

H_0 ditolak apabila $\chi^2_{Pearson} > \chi^2_{n-p-1}$ atau $P(\chi^2_{n-p-1} > \chi^2_{Pearson}) < \alpha$, sehingga dapat disimpulkan model tidak sesuai (Hosmer dan Lemeshow, 2000).

2.7. Least Absolute Shrinkage and Selection Operator

LASSO (*Least Absolute Shrinkage and Selection Operator*) merupakan metode yang dapat menyusutkan koefisien regresi menjadi nol dan sekaligus dapat melakukan pemilihan peubah prediktor (Tibshirani, 1996). Apabila pada regresi linier terdapat korelasi yang tinggi (multikolinieritas) antar peubah prediktor maka hasil pendugaan dengan menggunakan metode OLS menghasilkan penduga yang memiliki ragam besar.

Untuk menangani masalah tersebut dapat digunakan pemilihan peubah atau menerapkan LASSO. *Least Absolute Shrinkage and Selection Operator* (LASSO) diperkenalkan oleh Tibshirani (1996) sebagai metode yang dapat menyusutkan penduga parameter menjadi nol sekaligus melakukan pemilihan peubah.

LASSO dapat memminimumkan jumlah kuadrat galat pengamatan untuk jumlah nilai mutlak koefisien menjadi kurang dari konstanta. Dengan kata lain, LASSO membatasi penduga parameter kurang dari suatu konstanta, sehingga ada beberapa parameter yang bernilai nol. Nilai penduga LASSO dapat dituliskan sebagai

$$\hat{\beta}^L = \arg \min \sum_{i=1}^n (y_i - \beta_0 - \sum_{k=1}^p \beta_k x_{ik})^2 \quad (2.25)$$

dengan syarat $\sum_{k=1}^p |\hat{\beta}_k^L| \leq t$

di mana t merupakan parameter *tuning*. Apabila $t > \sum_{k=1}^p |\hat{\beta}_k^0|$ maka penduga LASSO akan sama dengan penduga OLS. Sehingga apabila $0 < t < \sum_{k=1}^p |\hat{\beta}_k^0|$, maka

$$\hat{\beta}^L = \arg \min \sum_{i=1}^n \left\{ \frac{1}{2} (y_i - \beta_0 - \sum_{k=1}^p \beta_k x_{ik})^2 + \lambda \sum_{k=1}^p |\hat{\beta}_k| \right\} \quad (2.26)$$

di mana:

y_i : nilai peubah respon ke- i

α : intersep

β_k : koefisien regresi peubah prediktor ke- k

x_{ik} : pengamatan ke- i peubah prediktor ke- k

i : $1, 2, \dots, n$; n adalah banyaknya pengamatan

k : $1, 2, \dots, p$; p adalah banyaknya peubah prediktor

Dengan memilih nilai t kecil, dapat menyebabkan beberapa koefisien menjadi nol. Nilai t harus dipilih sehingga dapat meminimumkan kesalahan prediksi. Metode pendugaan LASSO dilakukan secara bertahap dengan mencari nilai λ

$$\lambda = \frac{t}{\sum_{k=1}^p |\hat{\beta}_k|}$$

Nilai optimal dicari dengan proses *cross validation*. Lasso dapat membentuk beberapa koefisien menjadi sama dengan nol.

Selain dapat diterapkan pada OLS, LASSO juga dapat diterapkan pada regresi logistik. Penerapan LASSO pada regresi logistik dilakukan dengan penambahan pinalti LASSO pada fungsi $\log likelihood$ dalam proses penduga koefisien peubah prediktor (Hastie, dkk. 2008). Sehingga penduga regresi logistik dapat diperoleh dengan memaksimumkan fungsi $\log likelihood$.

Friedman dkk. (2010) dalam Hastie dkk. (2008) mengemukakan persamaan untuk penduga regresi logistik untuk peubah respon multinomial dengan penalti LASSO dituliskan sebagai

$$\hat{\beta}^L = \operatorname{argmax}\left\{ \ell(\hat{\beta}) - \lambda \sum_{j=1}^J \sum_{k=1}^p |\hat{\beta}_{jk}| \right\} \quad (2.27)$$

di mana $\ell(\hat{\beta})$ merupakan fungsi $\log likelihood$ dari pendugaan koefisien regresi logistik multinomial seperti di persamaan 2.16.

Untuk memperoleh hasil dari persamaan 2.27 diperlukan algoritma yang tepat. Salah satu algoritma yang dapat digunakan adalah algoritma *Cyclic Coordinate Descent* (CCD). Menurut Garcia-Margarinos, dkk (2009) kegunaan algoritma CCD adalah mengoptimasi nilai salah satu peubah sedangkan nilai peubah lainnya tetap (konstan). Ide algoritma CCD adalah membentuk batas atas turunan kedua dari tujuan pada selang nilai saat ini.

Tujuan algoritma ini adalah membentuk nilai parameter baru

$$\beta_k^{new} = \begin{cases} \beta_k - \Delta_k & \text{jika } \Delta v_k < -\Delta_k \\ \beta_k + \Delta v_k & \text{jika } -\Delta_k < \Delta v_k < \Delta_k \\ \beta_k + \Delta_k & \text{jika } \Delta_k < \Delta v_k \end{cases} \quad (2.28)$$

di mana:

$k : 1, 2, \dots, p$; p adalah banyaknya peubah prediktor

β_k : penduga parameter awal

Δ_k : hampiran penduga β_k

Δv_k : nilai penduga terbaru

Nilai interval $(\beta_k - \Delta_k, \beta_k + \Delta_k)$ berubah-ubah berdasarkan nilai penduga sebelumnya dan nilai Δv_k sebelumnya. Nilai Δv_k dapat diperoleh menggunakan persamaan:

$$\Delta v_k = -\frac{s_k(\beta) - \lambda s}{Q(\beta_k, \Delta_k)} \quad (2.29)$$

di mana :

$S_k(\beta)$: turunan pertama fungsi log *likelihood*

$Q(\beta_k, \Delta_k)$: turunan kedua fungsi log *likelihood*

$$S_k(\beta) = \frac{\partial L(\beta)}{\partial \beta_{jk}} = \sum_{i=1}^n x_{ki}(y_{ji} - \pi_{ji}) \quad (2.30)$$

$$Q(\beta_k, \Delta_k) = \frac{\partial^2 L(\beta)}{\partial \beta_{jk} \partial \beta_{jk}} = \sum_{i=1}^n x_{k'i} x_{ki} \pi_{ji} \pi_{j'i} \quad (2.31)$$

$$s = \frac{\beta_k}{|\beta_k|} \quad (2.32)$$

Menurut Zhang dan Oles (2001) pemilihan nilai awal Δ_k tidaklah berpengaruh. Nilai Δ_k dapat dimisalkan dalam $[1, \infty]$ atau dapat pula menggunakan rumus

$$\Delta_k = \max(2|\beta_k|, \frac{\Delta_k}{2}) \quad (2.33)$$

Algoritma CCD prosedurnya adalah sebagai berikut (Genkin, dkk. 2011):

1. Menentukan nilai-nilai awal $\beta_k = 0$ dan $\Delta_k = 1$ untuk $k=1, 2, \dots, p$; p adalah banyaknya peubah prediktor.
2. Menghitung Δv_k dengan ketentuan:
 - a. Apabila $\beta_k = 0$, maka $s = 1$, kemudian hitung Δv_k menggunakan persamaan (2.24). Apabila $\Delta v_k \leq 0$ maka $s = -1$ kemudian hitung Δv_k . Namun apabila $\Delta v_k \geq 0$ maka $\Delta v_k = 0$.
 - b. Apabila $\beta_k \neq 0$, maka $s = \frac{\beta_k}{|\beta_k|}$, kemudian hitung Δv_k menggunakan persamaan (2.29). Apabila $s(\beta_k + \Delta v_k) \leq 0$ maka $\Delta v_k = -\beta_k$.
3. Menghitung nilai Δ_k menggunakan persamaan (2.33).
4. Menghitung nilai $\beta_k^{L new}$ menggunakan persamaan (2.28).
5. Mengulangi langkah 2 sampai 4 hingga mencapai kondisi konvergen.

2.8. Interpretasi Parameter

Menurut Agresti (2007), regresi logistik diinterpretasikan menggunakan *odds ratio* (OR) yang nilainya sama dengan $\exp(\beta_{jk})$. *Odds ratio* adalah untuk mengetahui tingkat kecenderungan atau risiko hubungan peubah prediktor terhadap peubah respon. *Odds ratio* dapat dituliskan:

$$\widehat{OR} = \frac{\text{odds}_1}{\text{odds}_0} = \exp(\beta_{jk}) \quad (2.34)$$

di mana β_{jk} merupakan koefisien model kategori ke- j peubah prediktor ke- k . Menurut Kleinbaum dan Klein (2010), pada regresi logistik multinomial dengan peubah respon yang memiliki j kategori akan didapat kan sebanyak $j-1$ *odds ratio*. *Odds ratio* pertama adalah perbandingan peluang antara peubah respon kategori satu dengan peubah respon kategori pembanding dan *odds ratio* kedua adalah perbandingan peluang antara peubah respon kategori dua dengan peubah respon kategori pembanding.

Sedangkan untuk *odds ratio* antar logit dilakukan dengan melakukan perbandingan antar *odds ratio* pada peubah prediktor yang sama-sama berpengaruh. Sehingga dapat dituliskan sebagai berikut:

$$\widehat{OR} = \frac{\widehat{OR}_2}{\widehat{OR}_1} = \frac{\frac{odds_2}{odds_0}}{\frac{odds_1}{odds_0}} = \frac{odds_2}{odds_1} = \frac{exp(\beta_{2k})}{exp(\beta_{1k})} \quad (2.35)$$

2.9. Pengambilan Sampel

Dalam melakukan penelitian ini, digunakan metode pengambilan sampel untuk mewakili populasi yang ada. Menurut Cochran (1977), secara umum terdapat dua macam pengambilan sampel yaitu *probability sampling* dan *non probability sampling*. Pada *probability sampling*, setiap anggota populasi mempunyai kesempatan yang sama untuk diambil sebagai sampel, sedangkan non *probability sampling* dilakukan tanpa memberikan kesempatan yang sama pada anggota populasi untuk dijadikan sampel. Pada penelitian ini, pengambilan sampel dilakukan dengan *cluster probability sampling* di mana diambil sampel secara acak pada sebuah populasi yang homogen akan teteapi terbagi menjadi grup yang heterogen.

Ukuran sampel yang digunakan mengacu pada rumus Issac dan Michael dalam buku Sukardi (2004):

$$n = \frac{\chi^2 NP(1-P)}{\alpha^2(N-1)+\chi^2 P(1-P)} \quad (2.3)$$

di mana:

n : Ukuran sampel

N : Banyak populasi

P : Peluang memilih tidak melanjutkan (0.05)

α : Derajat kesalahan (0,05)

χ^2 : Nilai tabel *Chi-Square* dengan derajat bebas 1

Pengambilan secara acak dapat dilakukan dengan undian, menggunakan tabel, atau menggunakan pembangkit bilangan acak. Pengambilan sampel ini dilakukan apabila peubah yang diteliti relatif homogen. Langkah-langkah untuk melakukan *cluster sampling* menggunakan undian adalah

- Menentukan ukuran populasi yang diteliti.
- Menentukan ukuran sampel yang diinginkan.
- Menghitung banyak sampel dalam satu klaster.
- Menghitung banyak klaster yang dibutuhkan.
- Memberikan nomor untuk semua anggota klaster, kemudian dilakukan undian.
- Nomor klaster yang terambil dijadikan sampel dan terus dilakukan hingga mencapai banyak klaster yang dibutuhkan peneliti.

Metode pengambilan sampel dengan pembangkit bilangan acak dapat menggunakan kalkulator atau menggunakan komputer dengan bantuan *Microsoft Excel*.

2.10. Pendidikan

Menurut Undang-undang No. 20 Tahun 2003 mengenai sistem pendidikan nasional, pendidikan adalah usaha secara sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar para peserta didik dapat mengembangkan potensi dirinya secara aktif untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, dan keterampilan yang diperlukan untuk dirinya sendiri, masyarakat, bangsa, dan negara.

2.11. Sekolah Lanjutan SMA, SMK dan MA

Setelah kelulusan, siswa sekolah tingkat pertama (SMP dan MTs) seringkali meneruskan pendidikan ke tingkat selanjutnya tanpa pertimbangan jenis sekolah. Sebagian melanjutkan untuk mengikuti tren, faktor biaya, dan sebagainya. Beberapa lulusan SMP dan MTs tidak melanjutkan pendidikan ke jenjang yang lebih tinggi. Lagkah untuk memilih pendidikan tingkat menengah berikutnya (SMA, MA, SMK) harus difikirkan matang-matang karena prospek di lapangan kerja bisa dipengaruhi oleh pendidikan terakhir.

SMA adalah sekolah tingkat atas yang mempersiapkan para siswa untuk masuk perguruan tinggi umum, SMA Negeri berada di bawah naungan Dinas Pendidikan Provinsi.

Disamping itu ada MA (Madrasah Aliyah), jenis sekolah ini mempersiapkan para siswa untuk masuk ke perguruan tinggi islam maupun umum. Materi kurikulum berkisar 60% pengetahuan umum dan 40% ilmu agama.

Berikut ciri-ciri kurikulum SMA maupun MA:

1. Materi pembelajaran lebih mengarah pada teori daripada praktik
2. Tamatannya tidak dipersiapkan untuk langsung kerja
3. Tempat belajar lebih tertuju di sekolah

Sekolah tingkat ini sangat cocok untuk para siswa yang akan melanjutkan ke perguruan tinggi karena kajian teori yang mendalam.

Selanjutnya SMK adalah jenjang pendidikan sekolah tingkat atas yang mempersiapkan peserta didik untuk bekerja di lapangan secara langsung. Jenjang sekolah SMK dirancang untuk mencetak siswa yang siap kerja di lapangan. Berikut adalah ciri-ciri dari SMK:

1. Materi pembelajaran lebih mengarah pada praktek daripada teori.
2. Tamatannya dipersiapkan untuk siap kerja
3. Tempat belajar di sekolah dan di dunia kerja.

Namun lulusan SMA dan MA tidak menjadi jaminan untuk menjadi pengangguran dan SMK tidak menjamin mendapat peluang pekerjaan yang lebih baik. Semua tergantung usaha individu masing-masing dan kualitas pendidikan sangat mempengaruhi pola pikir siswa.

2.12. Faktor-faktor yang mempengaruhi pilihan sekolah.

Secara umum faktor-faktor yang mempengaruhi perencanaan pilihan SMA sederajat bagi siswa SMP dan MTs dibedakan menjadi dua yaitu faktor internal dan faktor eksternal. Faktor internal adalah nilai akademik siswa, tujuan jangka panjang, orang yang memberi pengaruh serta pertimbangan pribadi dan faktor eksternal terdiri dari pendidikan orang tua, pendapatan rata-rata per bulan orang tua, jumlah tanggungan dalam keluarga, dan pekerjaan orangtua. Menurut Pramudia (2006), faktor orangtua adalah faktor utama dalam mendorong anak untuk melanjutkan sekolah ke jenjang lebih tinggi.

1. Jenis sekolah yang sedang ditempuh.

Jenis dari sekolah yang sedang ditempuh oleh siswa berhubungan erat dengan sekolah lanjutan yang akan dipilih. Materi yang disampaikan di jenjang SMP atau MTs akan dilanjutkan di jenjang lebih tinggi dengan lebih rinci. Sebagai contoh, materi spesifik tentang agama yang diajarkan di Madrasah Aliyah (MA) sebelumnya juga ditempuh di Madrasah Tsanawiyah (MTs). Jenis sekolah yang terdapat di Kecamatan Munjungan adalah SMP dan MTs.

2. Jenis pekerjaan orangtua.

Pekerjaan merupakan sumber penghasilan yang digunakan untuk memenuhi kebutuhan keluarga. Terdapat berbagai jenis pekerjaan di Kecamatan Munjungan misal nelayan, petani, PNS, wiraswasta dan lain-lain. Nelayan dan petani merupakan pekerjaan yang banyak ditemui dan mendapatkan pendapatan dari pengelolaan hasil alam di laut bagi nelayan dan di darat bagi petani. Pegawai negeri sipil adalah satuan dinas yang bekerja di bawah naungan pemerintah, dalam hal ini gaji dan tunjangan didapatkan dari pemerintah. Wiraswasta merupakan seseorang yang menciptakan lapangan pekerjaan dan pendapatan didapatkan dari laba usaha. Selain pekerjaan yang telah disebutkan, masih ada pekerjaan lain seperti pekerja swasta, polisi, tentara dll.

3. Pendidikan terakhir orang tua.

Tingkat pendidikan adalah tahapan pendidikan yang ditetapkan berdasarkan tingkat perkembangan peserta didik, tujuan yang akan dicapai dan kemampuan yang dikembangkan. Orang tua yang telah menempuh jenjang pendidikan lebih tinggi cenderung memperhatikan dan memahami pentingnya pendidikan bagi anak-anaknya. Orangtua yang pernah mengenyam pendidikan lebih tinggi minimal SMA atau sederajat lebih

banyak mengetahui dunia pendidikan dalam kehidupan sehari-hari sehingga mampu memberikan saran pendidikan yang lebih baik.

4. Pendapatan rata-rata per bulan orangtua

Secara umum pendapatan adalah jumlah harta kekayaan awal periode ditambah perubahan penilaian yang tidak diakibatkan oleh perubahan modal atau hutang dari penghasilan orangtua rata-rata yang diperoleh orangtua dalam kurun waktu satu bulan. Siswa dengan sosial ekonomi keluarga yang baik akan mendapatkan fasilitas pendidikan yang baik .

5. Banyak tanggungan dalam keluarga

Banyak tanggungan dalam keluarga berhubungan erat dengan pendapatan orangtua. Meskipun pendapatan orang tua tinggi tetapi banyak tanggungan keluarga banyak, dapat berakibat berkurangnya biaya untuk sarana yang bisa diperoleh untuk anak.Hal tersebut dikarenakan orangtua membutuhkan biaya untuk kebutuhan lebih dari satu anak.

6. Nilai akademik

Belajar adalah proses yang dilakukan individu untuk memperoleh perubahan sikap, tingkah laku dan pengetahuan baru. Prestasi adalah hasil yang telah dikerjakan dan dicapai oleh setiap individu. Sedangkan prestasi akademik merupakan penilaian terhadap hasil dari belajar siswa untuk mengetahui sejauh mana siswa telah mencapai sasaran belajar. Prestasi individu bisa dilihat dari nilai ujian. Dalam penelitian ini nilai akademik yang diambil adalah nilai akhir semester 2 kelas VIII karena siswa kelas IX baru melaksanakan semester pertama.

7. Tujuan jangka panjang.

Setiap siswa harus memiliki tujuan dalam memilih pendidikannya. Tujuan jangka panjang dibutuhkan untuk mengetahui apa yang dilakukan siswa setelah mendapatkan bekal dari SLTA yang dipilih. Bagi siswa yang melanjutkan ke SMK, mereka akan dibekali kemampuan praktis untuk langsung bekerja akan tetapi tidak menutup kemungkinan siswa yang memilih SMA atau MA juga memilih bekerja setelah lulus. Menurut Djaali (2006), seseorang yang merasa lebih cekatan dalam bekerja akan berharap cepat mendapatkan timbalbalik. Alasan lain dalam pemilihan sekolah salah tentunya mendapatkan pengalaman dan pengetahuan tinggi sehingga banyak dari siswa SLTA memilih untuk melanjutkan pendidikan yang lebih tinggi.

8. Orang yang mempengaruhi

Siswa mendapatkan dorongan dari berbagai pihak dalam menentukan pilihan sekolah. Orangtua atau keluarga menjadi salah satu pihak yang sangat berpengaruh dalam menentukan pilihan sekolah anaknya. Selain itu, fakta menunjukkan bahwa siswa juga mendaftar pilihan sekolah lanjutan karena memiliki pemikiran yang sama dengan temannya karena siswa saling berdiskusi satu sama lain. Sekolah melalui guru dan bimbingan

konseling juga selalu memberikan pertimbangan kepada siswa tentang sekolah lanjutan yang sebaiknya dipilih. Dari berbagai pihak yang memberi masukan, siswa dapat dengan mudah menentukan sekolah yang akan diambil. Akan tetapi, tidak menutup kemungkinan siswa tersebut lebih memilih sekolah berdasarkan keinginannya sendiri dibandingkan keinginan orangtua.

9. Hal penting yang menjadi pertimbangan siswa

Banyak sekali pertimbangan siswa untuk masuk SLTA. Pertimbangan perlu dilakukan oleh siswa untuk membantu memilih sekolah lanjutan. Biaya menjadi salah satu alasan yang berhubungan erat dengan kemampuan orangtua menyekolahkan anaknya. Biaya yang mahal bisa mengurungkan niat siswa yang berasal dari keluarga menengah kebawah. Fasilitas juga menjadi pertimbangan dalam memilih sekolah, dengan fasilitas yang bagus siswa menjadi tertarik untuk memilih sekolah. Reputasi sekolah menjadi patokan penting untuk memilih sekolah, reputasi yang bagus akan menarik siswa yang banyak. Beberapa siswa juga memilih untuk bersekolah di lokasi yang dekat dengan tempat tinggal atau belajar di sekolah yang sama dengan teman-temannya sekolah di SMP maupun MTs.

BAB III

METODE PENELITIAN

3.1. Sumber Data

Data merupakan data primer berasal dari angket atau survai kepada responden. Responden yang menjadi objek penelitian adalah siswa SMP dan MTs Negeri kelas IX di Kabupaten Trenggalek.

Peubah yang digunakan:

1. Peubah respon bersifat kategori yaitu sekolah terdiri dari 4 kategori (Tidak Melanjutkan = 0, SMA = 1, SMK = 2, MA= 3).
2. Peubah prediktor disajikan pada Tabel 3.1.

Tabel 3.1.Peubah Prediktor

Prediktor	Keterangan	Kategori	
X ₁	Nilai rata-rata kelas VIII	Kontinu	
X ₂	Jenis Sekolah	SMP	0
		MTs	1
X ₃	Pekerjaan Ayah	Petani	0
		Wiraswasta	1
		PNS	2
		Lain-lain	3
		Petani/Nelayan	0
X ₄	Pekerjaan Ibu	Wiraswasta	1
		PNS	2
		Ibu Rumah Tangga	3
		Lain-lain	4
		SD	0
X ₅	Pendidikan Ayah	SMP sederajat	1
		SMA sederajat	2
		Sarjana	3
		SD	0
X ₆	Pendidikan Ibu	SMP sederajat	1
		SMA sederajat	2
		Sarjana	3
		Tidak berpenghasilan	0
X ₇	Pendapatan rata-rata ayah tiap bulan	<2 juta	1
		2-5 juta	2
		>5 juta	3
X ₈	Pendapatan rata-rata	Tidak berpenghasilan	0

	ibu tiap bulan	<2 juta	1
		2-5 juta	2
		>5 juta	3
X ₉	Banyak Tanggungan Keluarga	1	0
		2	1
		3	2
		>3	3
X ₁₀	Tujuan Jangka Panjang	Melanjutkan Pendidikan Tinggi	0
		Bekerja	1
		Belum Merencanakan	2
X ₁₁	Hal penting yang dipertimbangkan	Biaya yang murah	0
		Banyak teman yang dikenal	1
		Fasilitas yang baik	2
		Lokasi sekolah dekat	3
		Popularitas Sekolah	4
X ₁₂	Pihak yang paling mempengaruhi pilihan	Diri sendiri	0
		Keluarga	1
		Teman	2
		Sekolah	3

3.2. Metode Penarikan Sampel

Populasi adalah siswa SMP dan MTs Negeri kelas IX di Kabupaten Trenggalek. Sampel diambil menggunakan teknik *cluster random sampling* karena populasi dibagi dalam kluster berupa kecamatan.

Unit penelitian adalah siswa kelas IX dan dilakukan perhitungan ukuran sampel untuk mengetahui banyak kelas yang dibutuhkan. Banyak kelas yang diambil menggunakan rumus 2.26:

$$n = \frac{(3,841)(276)(0,95)(0,05)}{0,05^2(276) + (3,841)(0,95)(0,05)} = 57,883 \approx 58$$

Banyak kecamatan di Kabupaten Trenggalek adalah 14 kecamatan. Untuk mendapatkan klaster yang dipilih, dilakukan perhitungan untuk mengetahui banyak kelas:

$$\text{Banyak kelas pada klaster} = \frac{\text{Banyak populasi kelas}}{\text{Banyak klaster}} = \frac{276}{14} = 19,7 \approx 20$$

$$\text{Banyak klaster yang dipilih} = \frac{\text{Banyak sampel}}{\text{Banyak kelas per klaster}} = \frac{58}{20} = 2,9 \approx 3$$

Dari hasil perhitungan sampel dengan *cluster probability sampling*. Secara acak dipilih tiga kecamatan sebagai klaster yaitu Kecamatan Trenggalek, Kecamatan Kampak dan Kecamatan Munjungan. Banyak sampel untuk kelas dan banyak siswa dalam kelas bisa dilihat pada Tabel 3.2. dan Tabel 3.3.

Tabel 3.2. Sampel Kelas

No.	Kecamatan	Sekolah	Sampel (Kelas)
1	Trenggalek	SMPN 1 Trenggalek	5
		SMPN 3 Trenggalek	5
		SMPN 5 Trenggalek	5
		MTsN 1 Trenggalek	5
2	Kampak	SMPN 1 Kampak	8
		SMPN 2 Kampak	2
		SMPN 3 Kampak	2
		MTsN 2 Trenggalek	7
3	Munjungan	SMPN 1 Munjungan	7
		SMPN 2 Munjungan	2
		SMPN 3 Munjungan	2
		MTsN 3 Trenggalek	8
Jumlah			58

Tabel 3.3. Banyak siswa tiap kelas

No.	Sekolah	Kelas	Banyak Siswa
1	SMPN 1 Trenggalek	D	32
		F	29
		G	32
		H	30
		I	29
2	SMPN 3 Trenggalek	A	32
		B	32
		C	30
		D	31
		E	31
3	SMPN 5 Trenggalek	B	33
		C	34
		D	34
		E	29
		F	25

4	MTsN 1 Trenggalek	A	36
		B	35
		C	34
		D	33
		E	35
5	SMPN 1 Kampak	A	30
		B	26
		C	27
		D	30
		E	29
		F	29
		G	28
		H	29
6	SMPN 2 Kampak	A	20
		B	18
7	SMPN 3 Kampak	A	17
		B	19
8	MTsN 2 Trenggalek	A	30
		B	25
		C	27
		D	22
		E	27
		F	25
		G	23
9	SMPN 1 Munjungan	A	26
		B	23
		C	16
		D	20
		E	21
		F	19
		G	20
10	SMPN 2 Munjungan	A	18
		B	18
11	SMPN 3 Munjungan	A	21
		B	22
12	MTsN 3 Trenggalek	A1	22
		A2	23
		A3	23
		A4	21

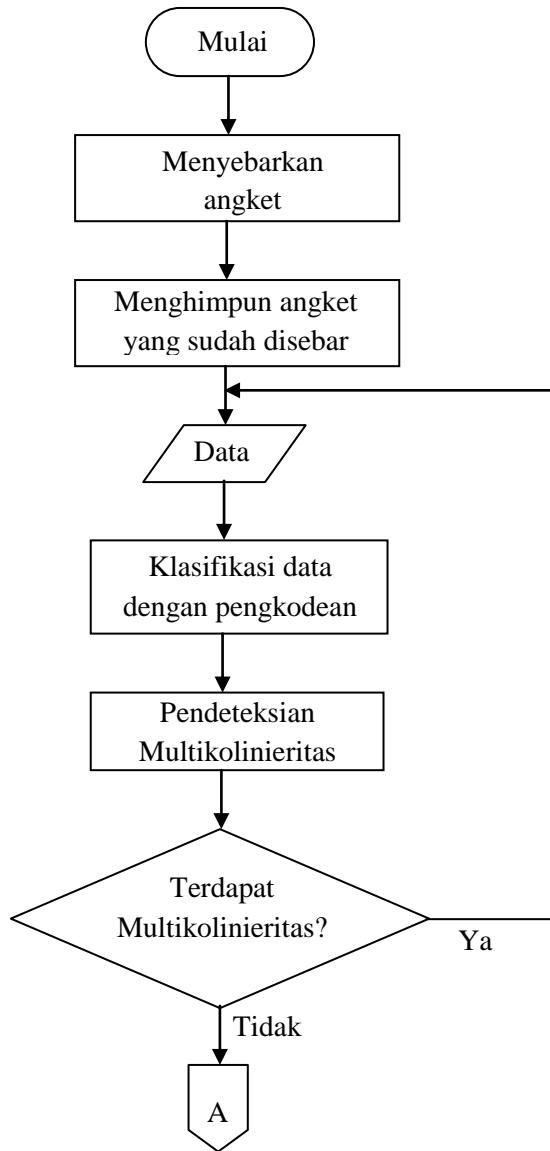
		C	27
		D	27
		E	26
		F	26
Jumlah			1536

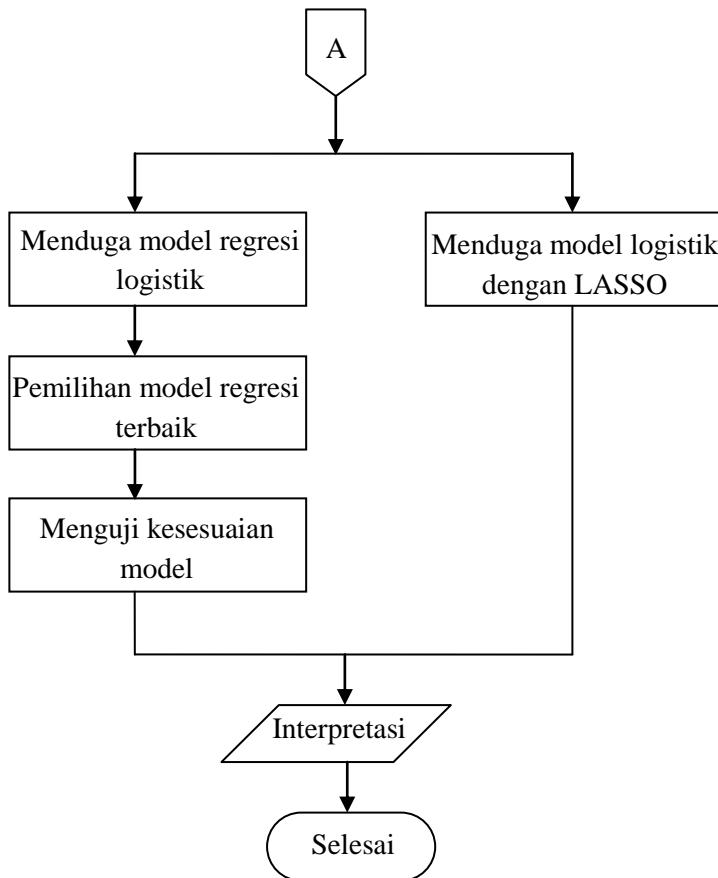
3.3. Metode Analisis

Langkah-langkah yang diterapkan pada data mencapai tujuan adalah sebagai berikut

1. Membuat angket, kemudian disebarluaskan kepada siswa SMP dan MTs Negeri di Kabupaten Trenggalek kelas IX yang terpilih menjadi sampel.
2. Menghimpun angket yang telah dibagikan di SMP dan MTs Negeri di Kabupaten Trenggalek.
3. Membentuk data menjadi data kategorik baik untuk peubah respon maupun peubah prediktor dengan melakukan pengkodean.
4. Mendeteksi adanya multikolinieritas dengan menggunakan persamaan (2.8) dengan ketentuan koefisien determinasi sebagai berikut
 - a. Apabila peubah respon model *auxiliary regression* berskala interval dan rasio menggunakan persamaan (2.9).
 - b. Apabila peubah respon model *auxiliary regression* berskala nominal dan ordinal menggunakan persamaan (2.10)
5. Regresi logistik multinomial
 - a. Menduga parameter regresi logistik multinomial
 - b. Melakukan pemilihan variabel berpengaruh dengan menggunakan statistik uji *Wald* pada persamaan (2.20)
 - c. Melakukan uji kesesuaian model dengan menggunakan statistik uji *Pearson residual* seperti pada persamaan (2.18)
6. Menduga parameter regresi logistik multinomial dengan LASSO menggunakan persamaan (2.28)
7. Interpretasi terhadap model yang telah diperoleh.

Pada gambar 3.1. berikut, ditunjukkan diagram alir penelitian menggunakan regresi logistik multinomial dengan metode LASSO.





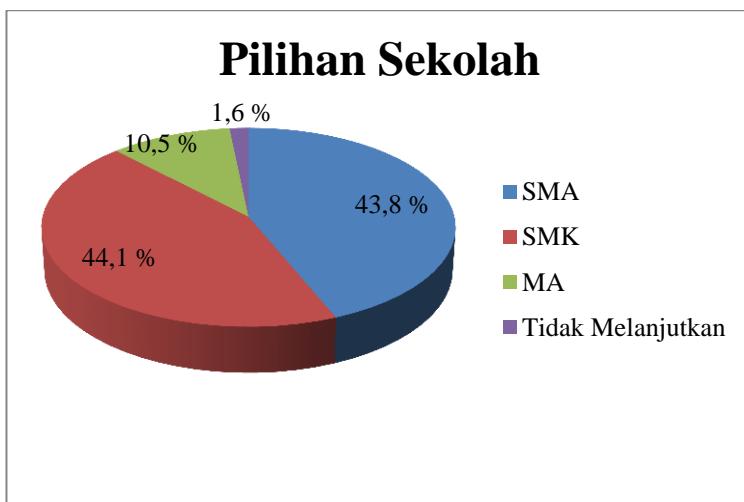
Gambar 3.1. Diagram Alir Penelitian

BAB IV

HASIL DAN PEMBAHASAN

4.1. Analisis Deskriptif

Analisis deskriptif digunakan untuk mengetahui gambaran umum mengenai data yang tersaji. Analisis deskriptif digunakan secara visual berupa diagram atau tabel untuk mempermudah penyampaian informasi. Penelitian ini menggunakan sampel siswa SMP dan MTs Negeri di Kabupaten Trenggalek. Banyak sampel yang digunakan adalah 1536 siswa. Diagram pilihan sekolah lanjutan setingkat SMA oleh siswa-siswi SMP dan MTs Negeri di Kabupaten Trenggalek tersaji dalam Gambar 4.1.



Gambar 4.1. Pilihan Sekolah Lanjutan Siswa SMP dan MTs Negeri di Kabupaten Trenggalek.

Terdapat 1536 siswa SMP dan MTs Negeri yang menjadi responden di Kabupaten Trenggalek. Berdasarkan Gambar 4.1, siswa yang memilih SMA sebesar 43,8% (672 siswa), SMK 44,1 % (677 siswa), MA 10,5 % (162 siswa), dan tidak melanjutkan sebesar 1,6 % (25 siswa). Angka paling banyak adalah siswa yang memilih SMK dibandingkan pilihan sekolah yang lain dan sedikit lebih banyak daripada yang memilih SMA, sedangkan siswa yang memilih MA lebih sedikit dibandingkan yang memilih dua sekolah lainnya, persentase terkecil adalah siswa yang tidak melanjutkan.

4.2. Pendektsian Multikolinieritas

Asumsi yang harus dipenuhi pada analisis regresi logistik multinomial adalah asumsi non multikolinieritas. Pendektsian multikolinieritas ini dapat dilakukan dengan melihat nilai VIF (*Variance Inflation Factors*) yang di dapatkan dari rumus 2.8. Apabila VIF bernilai lebih dari 10 maka, terdapat multikolinieritas antar peubah tersebut. Nilai VIF untuk masing-masing peubah prediktor disajikan pada Tabel 4.1.

Tabel 4.1. Nilai VIF untuk masing-masing Prediktor.

Peubah Prediktor	VIF
Nilai rata-rata kelas VIII(X1)	1,17
Jenis Sekolah (X2)	1,13
Pekerjaan Ayah (X3)	2,37
Pekerjaan Ibu (X4)	3,13
Pendidikan Ayah (X5)	2,39
Pendidikan Ibu (X6)	2,41
Pendapatan rata-rata ayah tiap bulan (X7)	1,45
Pendapatan rata-rata ibu tiap bulan (X8)	1,88
Banyak tanggungan keluarga (X9)	1,05
Tujuan Jangka Panjang (X10)	1,29
Hal penting yang dipertimbangkan (X11)	1,25
Pihak yang paling mempengaruhi pilihan (X12)	1,08

Dari Tabel 4.1 dapat dilihat bahwa tidak ada VIF yang bernilai lebih besar dari 10, dapat disimpulkan bahwa asumsi non multikolinieritas terpenuhi.

4.3. Regresi Logistik Multinomial

4.3.1. Pendugaan Parameter

Metode pendugaan parameter model yang digunakan merupakan metode *maximum likelihood* sesuai persamaan 2.17. Hasil nilai pendugaan parameter dapat dilihat pada Tabel 4.2.

Tabel 4.2. Penduga Parameter

Logit ke-1 (SMA)		Logit ke-2 (SMK)		Logit ke-3 (MA)	
Pendugaan Parameter		Pendugaan Parameter		Pendugaan Parameter	
β_{10}	55,465	β_{20}	58,784	β_{30}	45,071
β_{11}	0,240	β_{21}	0,175	β_{31}	0,181
$\beta_{12(0)}$	0,125	$\beta_{22(0)}$	-0,166	$\beta_{32(0)}$	-3,173
$\beta_{13(0)}$	-16,575	$\beta_{23(0)}$	-16,607	$\beta_{33(0)}$	-16,142
$\beta_{13(1)}$	-15,283	$\beta_{23(1)}$	-15,043	$\beta_{33(1)}$	-15,775
$\beta_{13(2)}$	-3,670	$\beta_{23(2)}$	-3,980	$\beta_{33(2)}$	-3,070
$\beta_{14(0)}$	-13,939	$\beta_{24(0)}$	-13,871	$\beta_{34(0)}$	-13,328
$\beta_{14(1)}$	-14,605	$\beta_{24(1)}$	-14,730	$\beta_{34(1)}$	-13,620
$\beta_{14(2)}$	-13,836	$\beta_{24(2)}$	-13,887	$\beta_{34(2)}$	-13,547
$\beta_{14(3)}$	-13,428	$\beta_{24(3)}$	-13,493	$\beta_{4(3)}$	-12,477
$\beta_{15(0)}$	-10,917	$\beta_{25(0)}$	-10,267	$\beta_{35(0)}$	-10,261
$\beta_{15(1)}$	-9,224	$\beta_{25(1)}$	-8,573	$\beta_{35(1)}$	-8,296
$\beta_{15(2)}$	-10,798	$\beta_{25(2)}$	-10,063	$\beta_{35(2)}$	-10,381
$\beta_{16(0)}$	-9,473	$\beta_{26(0)}$	-8,801	$\beta_{36(0)}$	-8,389
$\beta_{16(1)}$	-9,158	$\beta_{26(1)}$	-8,327	$\beta_{36(1)}$	-8,186
$\beta_{16(2)}$	-9,225	$\beta_{26(2)}$	-8,657	$\beta_{36(2)}$	-8,760
$\beta_{17(0)}$	-10,927	$\beta_{27(0)}$	-10,969	$\beta_{37(0)}$	-11,049
$\beta_{17(1)}$	-10,056	$\beta_{27(1)}$	-10,359	$\beta_{37(1)}$	-10,194
$\beta_{17(2)}$	-10,378	$\beta_{27(2)}$	-10,765	$\beta_{37(2)}$	-10,618
$\beta_{18(0)}$	-1,554	$\beta_{28(0)}$	-1,188	$\beta_{38(0)}$	15,060
$\beta_{18(1)}$	-1,237	$\beta_{28(1)}$	-,862	$\beta_{38(1)}$	15,178
$\beta_{18(2)}$	12,014	$\beta_{28(2)}$	11,869	$\beta_{38(2)}$	28,697
$\beta_{19(0)}$	4,024	$\beta_{29(0)}$	3,187	$\beta_{39(0)}$	2,940
$\beta_{19(1)}$	4,425	$\beta_{29(1)}$	3,609	$\beta_{39(1)}$	3,158
$\beta_{19(2)}$	3,429	$\beta_{29(2)}$	2,435	$\beta_{39(2)}$	2,572
$\beta_{110(0)}$	15,714	$\beta_{210(0)}$	14,671	$\beta_{310(0)}$	15,594
$\beta_{110(1)}$	-2,616	$\beta_{211(1)}$	-0,977	$\beta_{311(1)}$	-2,562
$\beta_{111(0)}$	-16,965	$\beta_{211(0)}$	-16,897	$\beta_{311(0)}$	-16,531

$\beta_{111(1)}$	-15,679	$\beta_{211(1)}$	-16,360	$\beta_{311(1)}$	-16,346
$\beta_{111(2)}$	-14,055	$\beta_{211(2)}$	-13,711	$\beta_{311(2)}$	-14,142
$\beta_{111(3)}$	-14,336	$\beta_{211(3)}$	-15,497	$\beta_{311(3)}$	-14,561
$\beta_{112(0)}$	0,257	$\beta_{212(0)}$	2,358	$\beta_{312(0)}$	-1,931
$\beta_{112(1)}$	-0,073	$\beta_{212(1)}$	1,887	$\beta_{312(1)}$	-2,501
$\beta_{112(2)}$	13,913	$\beta_{212(2)}$	15,505	$\beta_{312(2)}$	11,828

Pada pendugaan parameter regresi logistik multinomial diatas, referensi yang dipilih untuk setiap peubah prediktor kategorik adalah kategori terakhir untuk memudahkan interpretasi. Dalam penelitian ini akan terbentuk tiga model regresi logistik multinomial. Sebelum model regresi logistik terbentuk, dilakukan pengujian parameter model secara simultan dan parsial untuk mengetahui apakah peubah prediktor mempengaruhi peubah respon.

4.3.2. Pengujian Parameter Secara Simultan

Pengujian parameter secara simultan dilakukan dengan statistik uji G (*likelihood ratio test*) seperti pada persamaan 2.18 dengan hipotesis seperti pada bagian 2.5.1. Hasil analisis pengujian parameter secara simultan menggunakan *software SPSS* dapat dilihat pada tabel 4.3.

Tabel 4.3. Hasil Pengujian Parameter secara Simultan

Model	<i>Model Fitting Criteria</i>	Likelihood Ratio Tests		
		Khi - Kuadrat	df	<i>p-value</i>
<i>Intercept Only</i>	3134,812			
<i>Final</i>	2226,500	908,312	99	0,000

Tabel 4.3. menunjukkan bahwa *p-value* sangat kecil, sehingga H_0 ditolak. Dapat disimpulkan bahwa paling tidak terdapat satu peubah prediktor yang berpengaruh terhadap peubah respon.

4.3.3. Pengujian Parameter Secara Parsial

Pengujian parameter secara parsial menggunakan statistik uji *Wald* seperti pada persamaan 2.20. dengan hipotesis seperti pada bagian 2.5.2.

Hasil pengujian parameter secara lengkap dapat dilihat pada lampiran 9. Hasil analisis menggunakan SPSS secara ringkas dapat dilihat pada tabel 4.4.

Tabel 4.4. Hasil Pengujian Parameter secara Parsial

Logit	Peubah Prediktor	Kategori	β_{jk}	p-value
Logit 1 (SMA)	Nilai Rata-rata Semester 2 Kelas VIII		0,240	0,000
	Banyak tanggungan keluarga	0	4,024	0,000
		1	4,425	0,000
		2	3,429	0,001
Logit 2 (SMK)	Tujuan Jangka Panjang	1	-2,616	0,000
	Nilai Rata-rata Semester 2 Kelas VIII		0,175	0,006
	Banyak tanggungan keluarga	0	3,187	0,001
		1	3,609	0,000
		2	2,435	0,015
Logit 3 (MA)	Nilai Rata-rata Semester 2 Kelas VIII		0,181	0,007
	Jenis Sekolah	0	-3,173	0,000
	Banyak tanggungan keluarga	0	2,940	0,005
		1	3,158	0,003
		2	2,572	0,016
	Tujuan Jangka Panjang	1	-2,562	0,000

Tabel 4.4 menunjukkan parameter-parameter dengan $p\text{-value} < \alpha=0,05$, sehingga keputusan H_0 ditolak.

Dari persamaan 2.2 dihasilkan model regresi logistik multinomial:

- Model Regresi Logistik Multinomial untuk logit ke-1 (SMK):

$$\text{Logit}_1 = -0,240 x_1 + 4,024 x_{9(0)} + 4,425 x_{9(1)} - 3,429 x_{9(2)} - 2,616 x_{10(1)}$$

- Model Regresi Logistik Multinomial untuk logit ke-2 (MA):

$$\text{Logit}_2 = 0,175 x_1 + 3,187 x_{9(0)} + 3,609 x_{9(1)} + 2,435 x_{9(2)}$$

c. Model Regresi Logistik Multinomial untuk logit ke-3 (Tidak Melanjutkan):

$$\text{Logit}_3 = 0,181 x_1 - 3,173 x_{2(0)} + 2,940 x_{9(0)} + 3,158 x_{9(1)} + 2,572 x_{9(2)} - 2,562 x_{10(1)}$$

Sesuai persamaan 2.4, 2.5, 2.6, dan 2.7, peluang model regresi logistik multinomial yang terbentuk adalah:

1. Peluang model regresi logistik multinomial untuk Tidak Melanjutkan:

$$\pi_0 X_i = \frac{1}{1 + \exp(-0,240 x_1 + 4,024 x_{9(0)} + 4,425 x_{9(1)} - 3,429 x_{9(2)} - 2,616 x_{10(1)})} \\ + \exp(0,175 x_1 + 3,187 x_{9(0)} + 3,609 x_{9(1)} + 2,435 x_{9(2)}) \\ + \exp(0,181 x_1 - 3,173 x_{2(0)} + 2,940 x_{9(0)} + 3,158 x_{9(1)} + 2,572 x_{9(2)} \\ - 2,562 x_{10(1)})$$

2. Peluang model regresi logistik multinomial untuk SMA:

$$\pi_1 X_i = \frac{\exp(-0,240 x_1 + 4,024 x_{9(0)} + 4,425 x_{9(1)} - 3,429 x_{9(2)} - 2,616 x_{10(1)})}{1 + \exp(-0,240 x_1 + 4,024 x_{9(0)} + 4,425 x_{9(1)} - 3,429 x_{9(2)} - 2,616 x_{10(1)})} \\ + \exp(0,175 x_1 + 3,187 x_{9(0)} + 3,609 x_{9(1)} + 2,435 x_{9(2)}) \\ + \exp(0,181 x_1 - 3,173 x_{2(0)} + 2,940 x_{9(0)} + 3,158 x_{9(1)} + 2,572 x_{9(2)} \\ - 2,562 x_{10(1)})$$

3. Peluang model regresi logistik multinomial untuk SMK:

$$\pi_2 X_i = \frac{\exp(0,175 x_1 + 3,187 x_{9(0)} + 3,609 x_{9(1)} + 2,435 x_{9(2)})}{1 + \exp(-0,240 x_1 + 4,024 x_{9(0)} + 4,425 x_{9(1)} - 3,429 x_{9(2)} - 2,616 x_{10(1)})} \\ + \exp(0,175 x_1 + 3,187 x_{9(0)} + 3,609 x_{9(1)} + 2,435 x_{9(2)}) \\ + \exp(0,181 x_1 - 3,173 x_{2(0)} + 2,940 x_{9(0)} + 3,158 x_{9(1)} + 2,572 x_{9(2)} \\ - 2,562 x_{10(1)})$$

4. Peluang model regresi logistik multinomial untuk MA:

$$\pi_3 X_i = \frac{\exp(0,181 x_1 - 3,173 x_{2(0)} + 2,940 x_{9(0)} + 3,158 x_{9(1)} + 2,572 x_{9(2)} \\ - 2,562 x_{10(1)})}{1 + \exp(-0,240 x_1 + 4,024 x_{9(0)} + 4,425 x_{9(1)} - 3,429 x_{9(2)} - 2,616 x_{10(1)})} \\ + \exp(0,175 x_1 + 3,187 x_{9(0)} + 3,609 x_{9(1)} + 2,435 x_{9(2)}) \\ + \exp(0,181 x_1 - 3,173 x_{2(0)} + 2,940 x_{9(0)} + 3,158 x_{9(1)} + 2,572 x_{9(2)} \\ - 2,562 x_{10(1)})$$

4.3.4. Pengujian Kesesuaian Model

Pengujian kesesuaian model menggunakan statistik uji *Pearson* dilakukan untuk mengetahui apakah model sesuai untuk menggambarkan data berdasarkan hipotesis seperti pada bagian 2.6.

Hasil pengujian kesesuaian model disajikan pada Tabel 4.5.

Tabel 4.5. Hasil pengujian Kesesuaian Model.

	Khi-Kuadrat	db	p-value
Pearson	3539.434	4407	1.000

Tabel 4.5 menunjukkan $p\text{-value} > \alpha(0,05)$, sehingga H_0 diterima, model sesuai.

4.4. Regresi Logistik Multinomial dengan LASSO

Regresi Logistik Multinomial dengan LASSO dapat menyusutkan koefisien regresi menjadi nol dan sekaligus dapat melakukan pemilihan peubah prediktor. LASSO berguna pada data yang memiliki banyak peubah prediktor. Pemilihan model dilakukan dengan memilih nilai optimal λ melalui *cross validation*. Pada penelitian ini diperoleh nilai λ sebesar 0,02588615. Hasil pendugaan parameter regresi logistik multinomial menggunakan LASSO disajikan pada tabel 4.6

4.6. Penduga Parameter menggunakan LASSO

Logit ke-1 (SMA)		Logit ke-2 (SMK)		Logit ke-3 (MA)	
Pendugaan Parameter		Pendugaan Parameter		Pendugaan Parameter	
β_{10}	-2,053	β_{20}	1,308	β_{30}	0,746
β_{11}	0,038	β_{21}	0	β_{31}	0
$\beta_{12(0)}$	0,015	$\beta_{22(0)}$	0	$\beta_{32(0)}$	-2,028
$\beta_{13(0)}$	0	$\beta_{23(0)}$	0	$\beta_{33(0)}$	0,235
$\beta_{13(1)}$	0	$\beta_{23(1)}$	0,023	$\beta_{33(1)}$	0
$\beta_{13(2)}$	0	$\beta_{23(2)}$	-0,237	$\beta_{33(2)}$	0
$\beta_{14(0)}$	0	$\beta_{24(0)}$	0	$\beta_{34(0)}$	0
$\beta_{14(1)}$	0	$\beta_{24(1)}$	0	$\beta_{34(1)}$	0
$\beta_{14(2)}$	0,473	$\beta_{24(2)}$	0	$\beta_{34(2)}$	0
$\beta_{14(3)}$	0	$\beta_{24(3)}$	0	$\beta_{4(3)}$	0
$\beta_{15(0)}$	0	$\beta_{25(0)}$	0	$\beta_{35(0)}$	0
$\beta_{15(1)}$	0	$\beta_{25(1)}$	0	$\beta_{35(1)}$	0
$\beta_{15(2)}$	0	$\beta_{25(2)}$	0	$\beta_{35(2)}$	0
$\beta_{16(0)}$	0	$\beta_{26(0)}$	0	$\beta_{36(0)}$	0
$\beta_{16(1)}$	-0,098	$\beta_{26(1)}$	0	$\beta_{36(1)}$	0
$\beta_{16(2)}$	0	$\beta_{26(2)}$	0	$\beta_{36(2)}$	0
$\beta_{17(0)}$	0	$\beta_{27(0)}$	0	$\beta_{37(0)}$	0
$\beta_{17(1)}$	0	$\beta_{27(1)}$	0	$\beta_{37(1)}$	0
$\beta_{17(2)}$	0,025	$\beta_{27(2)}$	0	$\beta_{37(2)}$	0

$\beta_{18(0)}$	0	$\beta_{28(0)}$	0	$\beta_{38(0)}$	0
$\beta_{18(1)}$	0	$\beta_{28(1)}$	0	$\beta_{38(1)}$	0
$\beta_{18(2)}$	0,300	$\beta_{28(2)}$	0	$\beta_{38(2)}$	0
$\beta_{19(0)}$	0	$\beta_{29(0)}$	0	$\beta_{39(0)}$	0
$\beta_{19(1)}$	0	$\beta_{29(1)}$	0	$\beta_{39(1)}$	0
$\beta_{19(2)}$	0	$\beta_{29(2)}$	0	$\beta_{39(2)}$	0
$\beta_{110(0)}$	0,328	$\beta_{210(0)}$	-0,560	$\beta_{310(0)}$	0
$\beta_{110(1)}$	0	$\beta_{210(1)}$	1,256	$\beta_{310(1)}$	0
$\beta_{111(0)}$	0	$\beta_{211(0)}$	0	$\beta_{311(0)}$	0
$\beta_{111(1)}$	0	$\beta_{211(1)}$	0	$\beta_{311(1)}$	0
$\beta_{111(2)}$	0	$\beta_{211(2)}$	0,209	$\beta_{311(2)}$	0
$\beta_{111(3)}$	0	$\beta_{211(3)}$	-0,473	$\beta_{311(3)}$	0
$\beta_{112(0)}$	0	$\beta_{212(0)}$	0	$\beta_{312(0)}$	0
$\beta_{112(1)}$	0	$\beta_{212(1)}$	0	$\beta_{312(1)}$	0
$\beta_{112(2)}$	0	$\beta_{212(2)}$	0	$\beta_{312(2)}$	0

Pada tabel 4.6 terdapat banyak penduga parameter dengan koefisien yang bernilai nol atau dapat dikatakan bahwa koefisien regresi disusutkan menjadi nol. Pada pembentukan model regresi, peubah prediktor dengan penduga parameter bernilai nol tidak dimasukkan ke dalam model sehingga LASSO juga dapat melakukan seleksi peubah prediktor. Prediktor yang dimasukkan ke dalam model tersaji pada Tabel 4.7.

Tabel 4.7. Parameter Berpengaruh Menggunakan LASSO

Logit	Peubah Prediktor	Kategori	β_{jk}
Logit 1 (SMA)	Intersep		-2,053
	Nilai Rata-rata Semester 2 Kelas VIII		0,038
	Jenis Sekolah	0	0,015
	Pekerjaan Ibu	2	0,473
	Pendidikan Ibu	1	-2,616
	Pendapatan Rata-rata Ayah	2	0,025
	Pendapatan Rata-rata Ibu	2	0,300
	Tujuan Jangka Panjang	0	0,328

Logit 2 (SMK)	Intersep		1,308
	Pekerjaan Ayah	1	0,023
		2	-0,237
	Tujuan Jangka Panjang	0	-0,560
		1	1,256
	Hal Penting yang Dipertimbangkan	2	0,209
		3	-0,473
Logit 3 (MA)	Intersep		0,745
	Jenis Sekolah	0	-2,027
	Pekerjaan Ayah	0	0,235

Sedangkan untuk perbandingan peubah prediktor yang berpengaruh antara metode MLE dan LASSO dapat dilihat pada tabel 4.8. Untuk signifikansi parameter yang lebih lengkap dapat di lihat pada lampiran 12.

Tabel 4.8. Perbandingan Perbandingan Koefisien yang berpengaruh antara metode MLE dan Metode LASSO

Logit	Peubah Prediktor	Kategori	β_{jk} MLE	β_{jk} LASSO
Logit 1 (SMA)	Intersep			-2,053
	Nilai Rata-rata Semester 2 Kelas VIII		0,240	0,038
	Jenis Sekolah	0		0,015
	Pekerjaan Ibu	2		0,473
	Pendidikan Ibu	1		-2,616
	Pendapatan Rata-rata Ayah	2		0,025
	Pendapatan Rata-rata Ibu	2		0,300
	Banyak tanggungan keluarga	0	4,024	
		1	4,425	
		2	3,429	
	Tujuan Jangka Panjang	0		0,328
		1	-2,616	
Logit 2 (SMK)	Intersep			1,308
	Nilai Rata-rata Semester 2 Kelas VIII		0,175	
	Pekerjaan Ayah	1		0,023
		2		-0,237
	Banyak tanggungan keluarga	0	3,187	
		1	3,609	

		2	2,435	
Tujuan Jangka Panjang	0		-0,560	
	1		1,256	
Hal Penting yang Dipertimbangkan	2		0,209	
	3		-0,473	
Logit 3 (MA)	Intersep			0,745
	Nilai Rata-rata Semester 2 Kelas VIII		0,181	
	Jenis Sekolah	0	-3,173	-2,027
	Pekerjaan Ayah	0		0,235
	Banyak tanggungan keluarga	0	2,940	
		1	3,158	
		2	2,572	
	Tujuan Jangka Panjang	1	-2,562	

Berikut perbandingan AIC yang dari kedua metode:

Tabel 4.9. Nilai AIC

Metode	MLE	LASSO
AIC	2430,500	837.1146

Dari persamaan 2.2 dihasilkan model regresi logistik multinomial. Berikut model regresi logistik multinomial dengan metode LASSO:

a. Model Regresi Logistik Multinomial untuk logit ke-1 (SMA):

$$\text{Logit}_1 = -2,053 + 0,038 x_1 + 0,015 x_{2(0)} + 0,473 x_{4(2)} - \\ 2,616 x_{6(1)} + 0,025 x_{7(2)} + 0,300 x_{8(2)} + 0,328 x_{10(0)}$$

b. Model Regresi Logistik Multinomial untuk logit ke-2 (SMK):

$$\text{Logit}_2 = 1,308 + 0,023 x_{3(1)} - 0,237 x_{3(2)} - 0,560 x_{10(0)} + \\ 1,256 x_{10(1)} + 0,209 x_{11(2)} + 0,209 x_{11(3)}$$

c. Model Regresi Logistik Multinomial untuk logit ke-3 (MA):

$$\text{Logit}_3 = 0,745 - 2,027 x_{2(0)} + 0,235 x_{3(0)}$$

Sesuai persamaan 2.4, 2.5, 2.6, dan 2.7, peluang model regresi logistik multinomial menggunakan metode LASSO yang terbentuk adalah:

- Peluang model regresi logistik multinomial untuk Tidak Melanjutkan:

$$\pi_0 X_i = \frac{1}{1 + \exp(-2,053 + 0,038 x_1 + 0,015 x_{2(0)} + 0,473 x_{4(2)} - 2,616 x_{6(1)} \\ + 0,025 x_{7(2)} + 0,300 x_{8(2)} + 0,328 x_{10(0)}) + \exp(1,308 + 0,023 x_{3(1)} \\ - 0,237 x_{3(2)} - 0,560 x_{10(0)} + 1,256 x_{10(1)} + 0,209 x_{11(2)} + 0,209 x_{11(3)}) \\ \exp(0,745 - 2,027 x_{2(0)} + 0,235 x_{3(0)})}$$

- Peluang model regresi logistik multinomial untuk SMA:

$$\pi_1 X_i = \frac{\exp(-2,053 + 0,038 x_1 + 0,015 x_{2(0)} + 0,473 x_{4(2)} - 2,616 x_{6(1)} \\ + 0,025 x_{7(2)} + 0,300 x_{8(2)} + 0,328 x_{10(0)})}{1 + \exp(-2,053 + 0,038 x_1 + 0,015 x_{2(0)} + 0,473 x_{4(2)} - 2,616 x_{6(1)} \\ + 0,025 x_{7(2)} + 0,300 x_{8(2)} + 0,328 x_{10(0)}) + \exp(1,308 + 0,023 x_{3(1)} \\ - 0,237 x_{3(2)} - 0,560 x_{10(0)} + 1,256 x_{10(1)} + 0,209 x_{11(2)} + 0,209 x_{11(3)}) \\ \exp(0,745 - 2,027 x_{2(0)} + 0,235 x_{3(0)})}$$

- Peluang model regresi logistik multinomial untuk SMK:

$$\pi_2 X_i = \frac{\exp(1,308 + 0,023 x_{3(1)} - 0,237 x_{3(2)} - 0,560 x_{10(0)} + 1,256 x_{10(1)} \\ + 0,209 x_{11(2)} + 0,209 x_{11(3)})}{1 + \exp(-2,053 + 0,038 x_1 + 0,015 x_{2(0)} + 0,473 x_{4(2)} - 2,616 x_{6(1)} \\ + 0,025 x_{7(2)} + 0,300 x_{8(2)} + 0,328 x_{10(0)}) + \exp(1,308 + 0,023 x_{3(1)} \\ - 0,237 x_{3(2)} - 0,560 x_{10(0)} + 1,256 x_{10(1)} + 0,209 x_{11(2)} + 0,209 x_{11(3)}) \\ \exp(0,745 - 2,027 x_{2(0)} + 0,235 x_{3(0)})}$$

- Peluang model regresi logistik multinomial untuk MA:

$$\pi_3 X_i = \frac{\exp(0,745 - 2,027 x_{2(0)} + 0,235 x_{3(0)})}{1 + \exp(-2,053 + 0,038 x_1 + 0,015 x_{2(0)} + 0,473 x_{4(2)} - 2,616 x_{6(1)} \\ + 0,025 x_{7(2)} + 0,300 x_{8(2)} + 0,328 x_{10(0)}) + \exp(1,308 + 0,023 x_{3(1)} \\ - 0,237 x_{3(2)} - 0,560 x_{10(0)} + 1,256 x_{10(1)} + 0,209 x_{11(2)} + 0,209 x_{11(3)}) \\ \exp(0,745 - 2,027 x_{2(0)} + 0,235 x_{3(0)})}$$

4.5. Interpretasi Model

Setelah model didapatkan, langkah selanjutnya adalah melakukan interpretasi untuk memperoleh kesimpulan dari pendugaan parameter. Interpretasi regresi logistik multinomial menggunakan *odds ratio* untuk mengetahui kencenderungan hubungan suatu peubah prediktor terhadap peubah respon.

4.5.1 Interpretasi Model Regresi Logistik Multinomial

Odds ratio untuk masing-masing logit berdasarkan pendugaan parameter dengan metode *maximum likelihood* ditampilkan pada Tabel 4.10, 4.11, dan 4.12. Sedangkan untuk perbandingan *Odds Ratio* untuk kategori prediktor yang sama-sama berpengaruh masing-masing logit dapat dilihat pada Tabel 4.13.

Tabel 4.10. *Odds Ratio* Logit ke-1 (SMA) dengan Metode *Maximum Likelihood*

Logit	Peubah Prediktor	Kategori	β_{jk}	$\exp(\beta_{jk})$
Logit 1 (SMA)	Nilai Rata-rata Semester 2 Kelas VIII		0,240	1,271
	Banyak tanggungan keluarga	0	4,024	55,92
		1	4,425	83,535
		2	3,429	30,846
	Tujuan Jangka Panjang	1	-2,616	0,073

Tabel 4.10 menunjukkan *odds ratio* dengan Tidak Melanjutkan sebagai peubah respon pembanding, disimpulkan bahwa:

1. Siswa dengan nilai lebih tinggi memiliki kecenderungan 1,271 kali untuk memilih SMA dibandingkan tidak melanjutkan, sehingga semakin rendah nilai, siswa cenderung memilih tidak melanjutkan.
2. Siswa yang berasal dari keluarga dengan tanggungan satu anak memiliki kecenderungan 83,535 kali untuk memilih SMA dibanding tidak melanjutkan, daripada siswa yang berasal dari keluarga dengan tanggungan lebih dari tiga anak.
3. Siswa yang berasal dari keluarga dengan tanggungan dua anak memiliki kecenderungan 55,92 kali untuk memilih SMA dibanding tidak melanjutkan, daripada siswa yang berasal dari keluarga dengan tanggungan lebih dari tiga anak.
4. Siswa yang berasal dari keluarga dengan tanggungan tiga anak memiliki kecenderungan 30,846 kali untuk memilih SMA dibanding tidak melanjutkan, daripada siswa yang berasal dari keluarga dengan tanggungan lebih dari tiga anak.
5. Siswa yang berencana langsung bekerja memiliki kecenderungan 0,073 kali untuk memilih SMA dibanding tidak melanjutkan, daripada siswa yang belum merencanakan.

Tabel 4.11. *Odds Ratio* Logit ke-2 (SMK) dengan Metode *Maximum Likelihood*

Logit	Peubah Prediktor	Kategori	β_{jk}	$\exp(\beta_{jk})$
Logit 2 (SMK)	Nilai Rata-rata Semester 2 Kelas VIII		0,175	1,192

	Banyak tanggungan keluarga	0	3,187	24,216
		1	3,609	36,945
		2	2,435	11,410

Tabel 4.11 menunjukkan *odds ratio* dengan Tidak Melanjutkan sebagai peubah respon pembanding, disimpulkan bahwa:

1. Siswa dengan nilai lebih tinggi memiliki kecenderungan 1,192 untuk memilih SMK dibandingkan tidak melanjutkan, sehingga semakin rendah nilai, siswa cenderung memilih tidak melanjutkan.
2. Siswa yang berasal dari keluarga dengan tanggungan satu anak memiliki kecenderungan 24,216 kali untuk memilih SMK dibanding tidak melanjutkan, daripada siswa yang berasal dari keluarga dengan tanggungan lebih dari tiga anak.
3. Siswa yang berasal dari keluarga dengan tanggungan dua anak memiliki kecenderungan 36,945 kali untuk memilih SMK dibanding tidak melanjutkan, daripada siswa yang berasal dari keluarga dengan tanggungan lebih dari tiga anak.
4. Siswa yang berasal dari keluarga dengan tanggungan tiga anak memiliki kecenderungan 11,410 kali untuk memilih SMK dibanding tidak melanjutkan, daripada siswa yang berasal dari keluarga dengan tanggungan lebih dari tiga anak.

Tabel 4.12. *Odds Ratio Logit ke-3 (MA) dengan Metode Maximum Likelihood*

Logit	Peubah Prediktor	Kategori	β_{jk}	$\exp(\beta_{jk})$
Logit 3 (MA)	Nilai Rata-rata Semester 2 Kelas VIII		0,181	1,198
	Jenis Sekolah	0	-3,173	0,042
	Banyak tanggungan keluarga	0	2,940	18,925
		1	3,158	23,525
		2	2,572	13,093
	Tujuan Jangka Panjang	1	-2,562	0,077

Tabel 4.12 menunjukkan *odds ratio* dengan Tidak Melanjutkan sebagai peubah respon pembanding, disimpulkan bahwa:

1. Siswa dengan nilai lebih tinggi memiliki kecenderungan 1,198 untuk memilih MA dibandingkan tidak melanjutkan, sehingga semakin rendah nilai, siswa cenderung memilih tidak melanjutkan.

2. Siswa yang berasal dari SMP memiliki kecenderungan 0,042 kali untuk memilih MA dibanding tidak melanjutkan, daripada siswa yang berasal dari MTs.
3. Siswa yang berasal dari keluarga dengan tanggungan satu anak memiliki kecenderungan 18,925 kali untuk memilih MA dibanding tidak melanjutkan, daripada siswa yang berasal dari keluarga dengan tanggungan lebih dari tiga anak.
4. Siswa yang berasal dari keluarga dengan tanggungan dua anak memiliki kecenderungan 23,525 kali untuk memilih MA dibanding tidak melanjutkan, daripada siswa yang berasal dari keluarga dengan tanggungan lebih dari tiga anak.
5. Siswa yang berasal dari keluarga dengan tanggungan tiga anak memiliki kecenderungan 13,093 kali untuk memilih MA dibanding tidak melanjutkan, daripada siswa yang berasal dari keluarga dengan tanggungan lebih dari tiga anak.
6. Siswa yang berencana langsung bekerja memiliki kecenderungan 0,077 kali untuk memilih MA dibanding tidak melanjutkan, daripada siswa yang belum merencanakan.

Tabel 4.13. Perbandingan *Odds Ratio* untuk Prediktor yang Sama (MLE)

Peubah Prediktor	Kategori	<i>Odds Ratio</i> SMA	<i>Odds Ratio</i> SMK	<i>Odds Ratio</i> MA
Nilai Rata-rata Semester 2 Kelas VIII		1,271	1,192	1,198
Banyak tanggungan keluarga	0	55,92	24,216	18,925
	1	83,535	36,945	23,525
	2	30,846	11,410	13,093
Tujuan Jangka Panjang	1	0,073		0,077

Tabel 4.13 menunjukkan *odds ratio* dengan peubah prediktor yang sama-sama berpengaruh pada setidaknya dua logit yang berbeda sesuai persamaan 3.35. Dari tabel diatas, interpretasi yang dapat diperoleh adalah:

$$1. \quad \widehat{OR}_{SMA/SMK} = \frac{\widehat{OR}_{SMA}}{\widehat{OR}_{SMK}} = \frac{1,271}{1,192} = 1,066$$

$$\widehat{OR}_{SMA/MA} = \frac{\widehat{OR}_{SMA}}{\widehat{OR}_{MA}} = \frac{1,271}{1,198} = 1,061$$

$$\widehat{OR}_{MA/SMK} = \frac{\widehat{OR}_{MA}}{\widehat{OR}_{SMK}} = \frac{1,198}{1,192} = 1,005$$

Siswa dengan nilai rata-rata semester 2 kelas VIII lebih tinggi memiliki kecenderungan 1,066 kali untuk memilih SMA dibanding SMK. Siswa dengan nilai rata-rata semester 2 kelas VIII lebih tinggi memiliki kecenderungan 1,066 kali untuk memilih SMA dibanding MA. Nilai yang lebih tinggi juga memiliki kecenderungan 1,005 kali untuk memilih MA dibanding SMK. Sehingga Nilai yang lebih tinggi secara berurutan cenderung memilih SMA, MA kemudian SMK.

$$2. \quad \widehat{OR}_{SMA/SMK} = \frac{\widehat{OR}_{SMA}}{\widehat{OR}_{SMK}} = \frac{55,92}{24,216} = 2,309$$

$$\widehat{OR}_{SMA/MA} = \frac{\widehat{OR}_{SMA}}{\widehat{OR}_{MA}} = \frac{55,92}{18,925} = 2,954$$

$$\widehat{OR}_{SMK/MA} = \frac{\widehat{OR}_{SMK}}{\widehat{OR}_{MA}} = \frac{83,535}{18,925} = 1,279$$

Siswa dengan banyak tanggungan keluarga sebanyak satu anak memiliki kecenderungan 2,309 kali untuk memilih SMA daripada SMK, daripada yang memiliki tanggungan lebih dari tiga anak. Siswa dengan banyak tanggungan keluarga sebanyak satu anak memiliki kecenderungan 2,954 kali untuk memilih SMA daripada MA, daripada yang memiliki tanggungan lebih dari tiga anak. Sedangkan siswa dengan tanggungan satu anak memiliki kecenderungan 1,279 kali untuk memilih SMK dibanding MA, daripada yang memiliki tanggungan lebih dari tiga anak.

$$3. \quad \widehat{OR}_{SMA/SMK} = \frac{\widehat{OR}_{SMA}}{\widehat{OR}_{SMK}} = \frac{55,92}{36,945} = 2,261$$

$$\widehat{OR}_{SMA/SMK} = \frac{\widehat{OR}_{SMA}}{\widehat{OR}_{MA}} = \frac{55,92}{23,525} = 2,377$$

$$\widehat{OR}_{SMK/MA} = \frac{\widehat{OR}_{SMK}}{\widehat{OR}_{MA}} = \frac{36,945}{23,525} = 1,57$$

Siswa dengan banyak tanggungan keluarga sebanyak dua anak memiliki kecenderungan 2,261 kali untuk memilih SMA daripada SMK, daripada yang memiliki tanggungan lebih dari tiga anak. Siswa dengan banyak tanggungan keluarga sebanyak dua anak memiliki kecenderungan 2,377 kali untuk memilih SMA daripada MA, daripada yang memiliki tanggungan lebih dari tiga anak. Sedangkan siswa dengan tanggungan keluarga dua anak memiliki kecenderungan 1,57 kali untuk memilih SMK dibanding MA, daripada yang memiliki tanggungan lebih dari tiga anak.

$$4. \quad \widehat{OR}_{SMA/SMK} = \frac{\widehat{OR}_{SMA}}{\widehat{OR}_{SMK}} = \frac{30,846}{11,410} = 2,703$$

$$\widehat{OR}_{SMA/MA} = \frac{\widehat{OR}_{SMA}}{\widehat{OR}_{MA}} = \frac{30,846}{13,093} = 2,356$$

$$\widehat{OR}_{SMK/MA} = \frac{\widehat{OR}_{SMK}}{\widehat{OR}_{MA}} = \frac{11,410}{13,093} = 0,871$$

Siswa dengan banyak tanggungan keluarga sebanyak tiga anak memiliki kecenderungan 2,703 kali untuk memilih SMA daripada SMK, daripada yang memiliki tanggungan lebih dari tiga anak. Siswa dengan banyak tanggungan keluarga sebanyak tiga anak memiliki kecenderungan 2,356 kali untuk memilih SMA daripada MA, daripada yang memiliki tanggungan lebih dari tiga anak. Sedangkan siswa dengan tanggungan keluarga tiga anak memiliki kecenderungan 0,871 kali untuk memilih SMK dibanding MA, daripada yang memiliki tanggungan lebih dari tiga anak.

$$5. \quad \widehat{OR}_{SMA/MA} = \frac{\widehat{OR}_{SMA}}{\widehat{OR}_{MA}} = \frac{0,073}{0,077} = 0,948$$

Siswa yang merencakan untuk langsung bekerja memiliki kecenderungan 0,948 yang sama untuk memilih SMA dibanding MA, daripada yang belum merencanakan.

4.5.1 Interpretasi Model Regresi Logistik Multinomial dengan Metode LASSO

Odds ratio untuk masing-masing logit berdasarkan pendugaan parameter dengan metode LASSO ditampilkan pada Tabel 4.14, 4.15, dan 4.16. Sedangkan untuk perbandingan *Odds Ratio* untuk kategori prediktor yang sama-sama berpengaruh masing-masing logit dapat dilihat pada Tabel 4.17.

Tabel 4.14. *Odds Ratio* Logit ke-1 (SMA) dengan Metode LASSO

Logit	Peubah Prediktor	Kategori	β_{jk}	<i>Odds Ratio</i>
Logit 1 (SMA)	Intersep		-2,053	
	Nilai Rata-rata Semester 2 Kelas VIII		0,038	1,039
	Jenis Sekolah	0	0,015	1,015
	Pekerjaan Ibu	2	0,473	1,605
	Pendidikan Ibu	1	-2,616	0,906
	Pendapatan Rata-rata	2	0,025	1,026

	Ayah			
	Pendapatan Rata-rata Ibu	2	0,300	1,350
	Tujuan Jangka Panjang	0	0,328	1,388

Tabel 4.14 menunjukkan *odds ratio* dengan Tidak Melanjutkan sebagai peubah respon pembanding, disimpulkan bahwa:

1. Siswa dengan nilai lebih tinggi memiliki kecenderungan 1,039 untuk memilih SMA dibandingkan yang memilih tidak melanjutkan.
2. Siswa yang berasal dari SMP memiliki kecenderungan 1,015 kali untuk memilih SMA dibanding yang memilih tidak melanjutkan, daripada siswa yang berasal dari MTs.
3. Siswa yang ibunya bekerja sebagai PNS memiliki kecenderungan 1,605 kali untuk memilih SMA dibanding yang memilih tidak melanjutkan, daripada siswa yang ibunya bekerja di bidang lain-lain.
4. Siswa yang pendidikan terakhir ibunya SMP atau sederajat memiliki kecenderungan 0,906 kali untuk memilih SMA dibanding yang memilih tidak melanjutkan, daripada siswa yang ibunya berpendidikan terakhir Perguruan Tinggi.
5. Siswa yang ayahnya berpendapatan rata-rata tiap bulan antara 2 sampai 5 juta rupiah memiliki kecenderungan 1,026 kali untuk memilih SMA dibanding yang memilih tidak melanjutkan, daripada siswa yang ayahnya berpendapatan rata-rata tiap bulan lebih dari 5 juta rupiah.
6. Siswa yang ibunya berpendapatan rata-rata tiap bulan antara 2 sampai 5 juta rupiah memiliki kecenderungan 1,350 kali untuk memilih SMA dibanding yang memilih tidak melanjutkan, daripada siswa yang ibunya berpendapatan rata-rata tiap bulan lebih dari 5 juta rupiah.
7. Siswa yang ingin melanjutkan pendidikan tinggi, memiliki kecenderungan 1,388 kali untuk memilih SMA dibanding yang tidak melanjutkan, daripada yang belum merencanakan.

Tabel 4.15. *Odds Ratio* Logit ke-2 (SMK) dengan Metode LASSO

Logit	Peubah Prediktor	Kategori	β_{jk}	<i>Odds Ratio</i>
Logit 2 (SMK)	Intersep		1,308	
	Pekerjaan Ayah	1	0,023	1,024
		2	-0,237	0,789
	Tujuan Jangka Panjang	0	-0,560	0,571
		1	1,256	3,510
	Hal Penting yang Dipertimbangkan	2	0,209	1,233
		3	-0,473	0,623

Tabel 4.15 menunjukkan *odds ratio* dengan Tidak Melanjutkan sebagai peubah respon pembanding, disimpulkan bahwa:

1. Siswa yang ayahnya bekerja sebagai wiraswasta memiliki kecenderungan 1,024 kali untuk memilih SMK dibanding yang memilih tidak melanjutkan, daripada siswa yang ayahnya bekerja di bidang lain-lain.
2. Siswa yang ayahnya bekerja sebagai PNS memiliki kecenderungan 0,789 kali untuk memilih SMK dibanding yang memilih tidak melanjutkan, daripada siswa yang ayahnya bekerja di bidang lain-lain.
3. Siswa yang berencana melanjutkan pendidikan tinggi memiliki kecenderungan 0,571 kali untuk memilih SMK dibanding yang memilih tidak melanjutkan, daripada siswa yang belum merencanakan.
4. Siswa yang berencana langsung bekerja memiliki kecenderungan 3,510 kali untuk memilih SMK dibanding yang memilih tidak melanjutkan, daripada siswa yang belum merencanakan.
5. Siswa yang mempertimbangkan fasilitas yang baik sebagai hal paling penting memiliki kecenderungan 3,510 kali untuk memilih SMK dibanding yang memilih tidak melanjutkan, daripada siswa yang mempertimbangkan popularitas sekolah atau sekolah favorit.
6. Siswa yang mempertimbangkan lokasi sekolah dekat dengan rumah sebagai hal paling penting memiliki kecenderungan 0,623 kali untuk memilih SMK dibanding yang memilih tidak melanjutkan, daripada siswa yang mempertimbangkan popularitas sekolah atau sekolah favorit.

Tabel 4.16. *Odds Ratio* Logit ke-3 (MA) dengan Metode LASSO

Logit	Peubah Prediktor	Kategori	β_{jk}	<i>Odds Ratio</i>
Logit 3 (MA)	Intersep		0,745	
	Jenis Sekolah	0	-2,027	0,132
	Pekerjaan Ayah	0	0,235	1,265

Tabel 4.16 menunjukkan *odds ratio* dengan Tidak Melanjutkan sebagai peubah respon pembanding, disimpulkan bahwa:

1. Siswa yang berasal dari SMP memiliki kecenderungan 0,132 kali untuk memilih MTs dibanding yang memilih tidak melanjutkan, daripada siswa yang berasal dari MTs.
2. Siswa yang ayahnya bekerja sebagai petani atau nelayan memiliki kecenderungan 1,265 kali untuk memilih MA dibanding yang memilih tidak melanjutkan, daripada siswa yang ayahnya bekerja di bidang lain-lain.

Tabel 4.17. Perbandingan *Odds Ratio* untuk Prediktor yang Sama (LASSO)

Peubah Prediktor	Kategori	<i>Odds Ratio</i> SMA	<i>Odds Ratio</i> SMK	<i>Odds Ratio</i> MA
Jenis Sekolah	0	1,015		0,132
Tujuan Jangka Panjang	0	1,388	0,571	

Tabel 4.17 menunjukkan *odds ratio* dengan peubah prediktor yang sama-sama berpengaruh pada setidaknya dua logit yang berbeda sesuai persamaan 3.35. Dari tabel diatas, interpretasi yang dapat diperoleh adalah:

$$1. \quad \widehat{OR}_{SMA/MA} = \frac{\widehat{OR}_{SMA}}{\widehat{OR}_{MA}} = \frac{1,015}{0,132} = 7,689$$

Siswa yang berasal dari SMP memiliki kecenderungan 7,689 kali untuk memilih SMA dibanding MA, daripada siswa yang berasal dari MTs.

$$2. \quad \widehat{OR}_{SMA/SMK} = \frac{\widehat{OR}_{SMA}}{\widehat{OR}_{SMK}} = \frac{1,388}{0,571} = 2,431$$

Siswa yang berencana melanjutkan pendidikan tinggi memiliki kecenderungan 2,431 kali untuk memilih SMA dibanding SMK, daripada siswa yang belum merencanakan.

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil analisis dan pembahasan dapat diambil kesimpulan bahwa:

1. Menggunakan regresi logistik multinomial, dari dua belas peubah prediktor dalam penelitian ini, terdapat empat peubah prediktor yang berpengaruh terhadap pilihan sekolah di Kabupaten Trenggalek yaitu nilai rata-rata semester dua saat kelas delapan, jenis sekolah, banyak tanggungan keluarga, dan tujuan jangka panjang.
2. Menggunakan metode LASSO, terdapat sembilan peubah prediktor yang berpengaruh terhadap pilihan sekolah di Kabupaten Trenggalek yaitu nilai rata-rata semester dua saat kelas delapan, jenis sekolah, pekerjaan ayah, pekerjaan ibu, pendidikan ibu, pendapatan ayah, pendapatan ibu, tujuan jangka panjang dan hal yang dipertimbangkan. Hasil analisis didasarkan pada data primer dengan menyebarluaskan angket pada siswa SMP dan MTs Negeri di Kabupaten Trenggalek.
3. Semakin tinggi nilai siswa, siswa akan cenderung memilih SMA diikuti MA dan SMK. Semakin banyak tanggungan keluarga siswa, semakin besar kecenderungan untuk tidak melanjutkan sekolah. Siswa yang ingin melanjutkan pendidikan tinggi cenderung memilih SMA dan siswa yang ingin langsung bekerja cenderung memilih SMK. Siswa yang ingin melanjutkan ke MA cenderung berasal dari siswa MTs. Semakin tinggi pendidikan dan pendapatan orangtua, siswa cenderung memilih SMA, sedangkan semakin rendah pendidikan terakhir orangtua siswa cenderung memilih tidak melanjutkan.

5.1. Saran

1. Berdasarkan penelitian ini, saran yang diberikan SMP dan MTs Negeri di Kabupaten Trenggalek adalah meningkatkan sistem pembelajaran yang tinggi karena nilai berpengaruh pada pilihan siswa. Selain itu, sekolah diharapkan memberikan konseling kepada siswa tentang pilihan sekolah agar siswa tepat dalam memilih

sekolah dengan mempertimbangkan tujuan dan kondisi keluarga siswa.

2. Untuk penelitian selanjutnya diharapkan dapat membandingkan metode LASSO dengan metode lainnya seperti metode *ridge* atau *elastic net* dalam melakukan seleksi peubah prediktor dengan menyesuaikan struktur data dan permasalahan yang dihadapi.

DAFTAR PUSTAKA

- Agresti, A. 2007. *An Introduction to Categorical Data Analysis, Second Edition.* John Wiley and Sons, Inc. New York.
- Cochran, W. G. 1977. Sampling Techniques, Third Edition. John Wiley and Sons. New York
- Djaati. 2006. *Psikologi Pendidikan.* BumiAksara. Jakarta
- Garcia-Margarinos, M.G., Antoniadis, A., Cao, R. dan Gonzalez-Manteiga, W. 2009. Lasso Logistic Regression, GSoft and the Cyclic Coordinate Descent Algorithm. *Application to Gene Expression Data. Statistical Application in Genetical and Molecular Biology*, Vol. 9 No. 1.
- Genkin, A., Lewis, D. D., dan Madigan, D. 2011. Sparse Logistic Regression for Text Categorization. *Pattern Recognition Letters*. Vol. 32. No. 2.
- Gujarati, D. N. 2004. *Basic Econometrics, Fourth Edition.* McGraw-Hill Companies.
- Hastie, T., Tibshirani, R. dan Friedman, J. 2008. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction , Second Edition.* John Springer. New York.
- Hosmer, D.W. dan Lemeshow, S. 2000. *Applied Logistic Regression.* John Wiley and Sons, Inc. New York
- Kementrian Pendidikan dan Kebudayaan. *Proyeksi Siswa Tingkat Nasional Tahun 2012/2013 - 2020/2021.* 2013. Pusat Data dan Statistik Pendidikan.
- Kleinbaum, D. G. dan Klein, M. 2010. *Logistic Regression: A Self Learning Text, Third Edition.* Springer. New York

- Pramudia, R. 2006. Orientasi Pendidikan. *Jurnal Pendidikan Luar Sekolah Universitas Pendidikan Indonesia*, 1(3), 1-2.
- Sukardi. 2004. *Metodologi Penelitian Pendidikan, Kompetensi dan Prakteknya*. BumiAksara. Jakarta
- Tibshirani, R. 1996. Regression Shrinkage and Selection via The Lasso. *Royal Statistical Society*. Vol. 58. No. 1.
- Zhang, T. dan Oles, F.J. Text Categorization Based on Regularized Linear Classification Methods. *Information Retrieval*, Vol. 4. No. 1.

Lampiran 1. Angket



KEMENTERIAN RISET, TEKNOLOGI DAN PENDIDIKAN TINGGI
UNIVERSITAS BRAWIJAYA FAKULTAS MIPA
JURUSAN MATEMATIKA

Jl. Veteran, Malang 65145, Jawa Timur, Indonesia, Telp-fax : +62-341-571142
<http://matematika.ub.ac.id>, e-mail: jurmatub@ub.ac.id

No. Responden

Kepada Yth.
Saudara/i
Responden Penelitian
di tempat

Dengan hormat,

Peneliti adalah mahasiswa Program Studi Statistika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Brawijaya.

Nama : Hafid Wahyu Ramadhan

NIM : 135090500111008

Saat ini peneliti sedang melakukan penelitian skripsi yang berjudul "**Pendugaan Parameter Regresi Logistik Multinomial dengan Metode Least Absolute Shrinkage and Selection Operator (Lasso)** Studi Kasus Pilihan sekolah sederajat SMA oleh Siswa SMP dan MTs Negeri Kabupaten Trenggalek". Untuk itu peneliti melakukan penyebaran angket guna mengumpulkan data.

Sehubungan dengan hal ini, peneliti memohon untuk kesediaan saudara/i meluangkan waktu dan berkenan memberikan jawaban pada angket. Jawaban saudara/i merupakan informasi yang sangat berarti, oleh karena itu kelengkapan pengisian angket dan kejujuran dalam menjawab pertanyaan-pertanyaan sangat peneliti harapkan.

Perlu diketahui, angket yang Saudara/i isi, hanya akan digunakan untuk penelitian, dan tidak akan digunakan untuk kepentingan lainnya.

Akhir kata, peneliti mengucapkan terima kasih kepada Saudara/i telah bersedia membantu pengisian angket ini.

Trenggalek, 5 September 2017
Peneliti

Hafid Wahyu Ramadhan

Lampiran 1. Lanjutan

ANGKET

Sehubungan dengan penelitian yang akan saya lakukan mohon diisi angket ini dengan sebenar-benarnya.

Mohon lengkapi biodata berikut:

1. Nama : _____
2. Asal Sekolah : _____
3. Kelas : _____
4. Jenis Kelamin : L/P (Coret yang tidak Perlu)
5. Nilai Rata-rata Semester II saat Kelas VIII: _____

Jawab pertanyaan di bawah ini dengan memberi tanda centang (✓) pada kotak yang telah disediakan. Centanglah hanya pada satu kotak yang disediakan untuk masing-masing pertanyaan.

1. Pekerjaan Ayah:
Petani/Nelayan Wiraswasta PNS Lain-lain, _____
2. Pekerjaan Ibu:
Petani/Nelayan Wiraswasta PNS Ibu Rumah Tangga
Lain-lain, _____
3. Pendidikan Terakhir Ayah:
SD SMP Sederajat SMA Sederajat Perguruan Tinggi
4. Pendidikan Terakhir Ibu:
SD SMP Sederajat SMA Sederajat Perguruan Tinggi
5. Pendapatan rata-rata ayah tiap bulan:
Tidak berpenghasilan < 2 juta 2 – 5 Juta > 5 juta
6. Pendapatan rata-rata ibu tiap bulan:
Tidak berpenghasilan < 2 juta 2 – 5 Juta > 5 juta
7. Banyak tanggungan dalam keluarga (termasuk Anda):
1 anak 2 anak 3 anak >3 anak
8. Pilihan melanjutkan:
SMA SMK MA Tidak melanjutkan
9. Tujuan kedepan Anda memilih pilihan nomor 8 adalah untuk:
Melanjutkan pendidikan tinggi (Universitas, Sekolah Tinggi, Akademi dst.)
Langsung bekerja
Belum Merencanakan
10. Hal yang paling penting dalam memilih melanjutkan ke SMA/MA/SMK menurut Anda adalah:
Biaya yang murah Banyak teman yang dikenal
Fasilitas yang baik Lokasi sekolah dekat dengan rumah
Popularitas Sekolah / Sekolah Favorit
11. Pihak yang paling mempengaruhi Anda dalam memilih adalah:
Tidak Ada/Saya Sendiri Keluarga
Teman Sekolah

Terimakasih telah mengisi angket untuk membantu penelitian yang saya lakukan.

LAMPIRAN

Lampiran 2. Tabulasi Hasil Penyebaran Angket Pilihan Sekolah Siswa SMP dan MTS Negeri di Kabupaten Trenggalek

Peubah	Kategori	Peubah Respon				Total
		SMA	SMK	MA	Tidak Melanjutkan	
Nilai Rata-rata (X1)	Kontinu	672	677	162	25	1536
Jenis Sekolah (X2)	SMP	469	485	17	18	989
	MTs	203	192	145	7	547
Jumlah		672	677	162	25	1536
Pekerjaan Ayah (X3)	Petani	226	290	99	21	636
	Wiraswasta	235	271	29	4	539
	PNS	105	29	12	0	146
	Lain-lain	106	87	22	0	215
Jumlah		672	677	162	25	1536
Pekerjaan Ibu (X4)	Petani/Nelayan	71	98	23	8	200
	Wiraswasta	132	121	20	4	277
	PNS	86	14	5	0	105
	Ibu Rumah Tangga	343	412	109	13	877
	Lain-lain	40	32	5	0	77
Jumlah		672	677	162	25	1536
Pendidikan Ayah (X5)	SD	176	263	65	20	524
	SMP sederajat	126	162	51	2	341
	SMA sederajat	228	224	35	3	490
	Sarjana	142	28	11	0	181
Jumlah		672	677	162	25	1536

Lampiran 2. Lanjutan

Peubah	Kategori	Peubah Respon				Total
		SMA	SMK	MA	Tidak Melanjutkan	
Pendidikan Ibu (X6)	SD	138	204	53	17	412
	SMP sederajat	159	232	57	5	453
	SMA sederajat	232	215	43	3	493
	Sarjana	143	26	9	0	178
Jumlah		672	677	162	25	1536
Pendapatan rata-rata ayah tiap bulan (X7)	Tidak berpenghasilan	27	45	8	5	85
	<2 juta	396	493	122	19	1030
	2-5 juta	217	125	29	1	372
	>5 juta	32	14	3	0	49
Jumlah		672	677	162	25	1536
Pendapatan rata-rata ibu tiap bulan (X8)	Tidak berpenghasilan	261	314	89	11	675
	<2 juta	273	318	60	14	665
	2-5 juta	126	37	13	0	176
	>5 juta	12	8	0	0	20
Jumlah		672	677	162	25	1536
Banyak Tanggungan (X9)	1	162	191	47	6	406
	2	356	357	72	8	793
	3	135	100	34	7	276
	>3	19	29	9	4	61
Jumlah		672	677	162	25	1536

Lampiran 2. Lanjutan

Peubah	Kategori	Peubah Respon				Total
		SMA	SMK	MA	Tidak Melanjutkan	
Tujuan Jangka Panjang (X10)	Melanjutkan Pendidikan Tinggi	491	212	95	0	798
	Bekerja	39	260	9	14	322
	Belum Merencanakan	142	205	58	11	416
Jumlah		672	677	162	25	1536
Hal penting yang dipertimbangkan (X11)	Biaya yang murah	19	38	12	9	78
	Banyak teman yang dikenal	34	23	8	6	71
	Fasilitas yang baik	323	416	84	8	831
	Lokasi sekolah dekat	59	28	16	2	105
	Popularitas Sekolah	237	172	42	0	451
Jumlah		672	677	162	25	1536
Pihak yang paling mempengaruhi pilihan (X12)	Diri sendiri	286	329	91	11	717
	Keluarga	357	324	64	13	758
	Teman	25	22	6	0	53
	Sekolah	4	2	1	1	8
Jumlah		672	677	162	25	1536

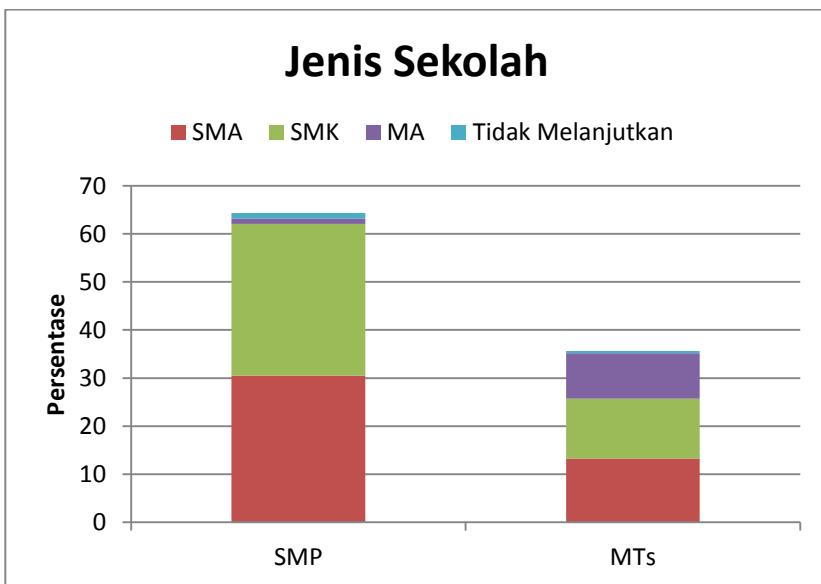
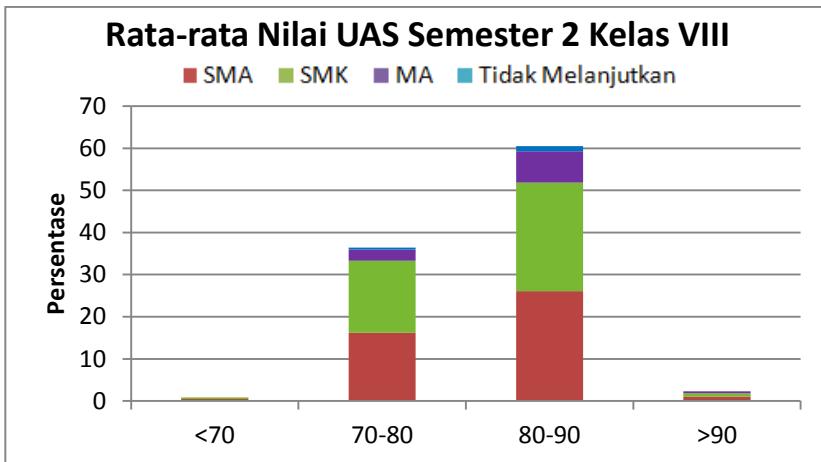
Lampiran 3. Data Pilihan Sekolah Siswa SMP dan MTs Negeri di Kabupaten Trenggalek.

No	Nama	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂
1	Imroatus Sholihah	0	86	0	1	3	2	1	1	0	2	0	4	1
2	Galih Pudha P.	0	85	0	3	3	3	3	3	0	1	0	4	0
3	Sisca Putri P.	0	87.31	0	2	3	2	2	2	0	0	0	4	0
4	Nuansa Bintang F.	0	86.5	0	2	2	3	3	2	2	1	0	2	1
5	Asista Arum J.	0	89.13	0	2	2	3	3	2	2	0	0	4	0
6	Intan Surya S.	0	90	0	1	3	2	1	3	0	1	0	4	1
7	Cahya Nursila A.	0	88.54	0	2	2	3	3	2	2	1	0	4	1
8	Anifa Qoriyah	0	88.59	0	0	3	0	0	1	0	2	0	4	0
9	Dito Pandu D.	0	87.5	0	3	3	2	2	1	0	1	0	4	0
10	Sintya Abila	0	88.18	0	0	3	1	1	1	0	0	0	4	0
11	Dia Rihadatul A.	0	88.16	0	1	3	1	2	1	0	0	0	4	0
12	Wahyu Dwi D.	0	86.81	0	1	3	2	2	1	0	1	0	4	1
13	Tsabudin A. N. N.	0	83	0	2	3	3	2	1	1	3	0	4	0
14	Mohamad R. A. H.	0	80	0	1	2	3	3	2	2	3	0	4	0
15	Iqbal Britanza	0	86	0	2	2	3	3	2	2	1	0	4	0
16	Tri Ananda Yoga U.	0	80	0	3	1	3	2	0	1	1	0	2	1
17	Farhan Pandu P.	0	85	0	2	2	3	3	2	2	0	2	4	0
18	Diara Isma A.	0	86	0	2	3	3	2	2	0	1	0	2	1

Lampiran 3. Lanjutan

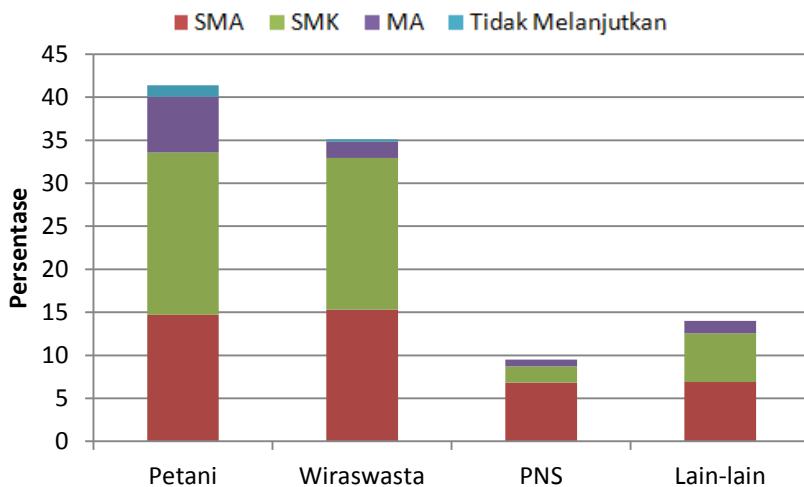
No	Nama	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	Y
19	Reisi Camar C.	1	85	0	2	1	3	2	2	1	2	0	2	0
20	Breian Danesti	0	85	0	2	2	2	3	2	2	2	0	2	1
21	Riski Amalina P.	0	86.86	0	1	3	2	2	3	0	2	0	2	1
22	Risa Kurniasari	0	86.7	0	1	3	2	2	1	0	1	0	4	1
23	Ainur Laili Z.	1	88.04	0	1	1	2	2	2	2	1	1	2	0
24	Sekar Arum P.	0	87.38	0	3	1	2	3	1	1	0	0	4	0
25	Hidayatus S.	1	80	0	3	3	3	2	2	0	1	2	2	0
26	Sandra Puja A.	0	86.7	0	1	1	2	3	3	3	1	0	2	1
27	Riska Ilham F.	0	80	0	1	3	3	3	2	0	1	2	4	0
28	Rosalinda D. P.	0	90	0	3	1	2	2	2	1	1	0	4	0
29	Vincent R. G. A.	0	82	0	1	1	2	2	2	2	0	2	2	0
30	Wirayuda	0	83	0	1	1	3	3	3	3	1	2	4	1
31	Reyegiuse E. N.	0	86.5	0	0	1	2	2	2	2	0	0	2	1
32	Aprilia Lutfi Z.	2	85	0	1	3	2	2	1	0	1	0	4	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1535	Akmal Maulana	0	90	1	2	3	2	2	2	0	2	1	2	1
1536	Mudatul Askola	1	75	1	0	0	0	0	1	1	0	0	2	1

Lampiran 4. Grafik Perbandingan Pilihan Perdasarkan Masing-Masing Peubah.

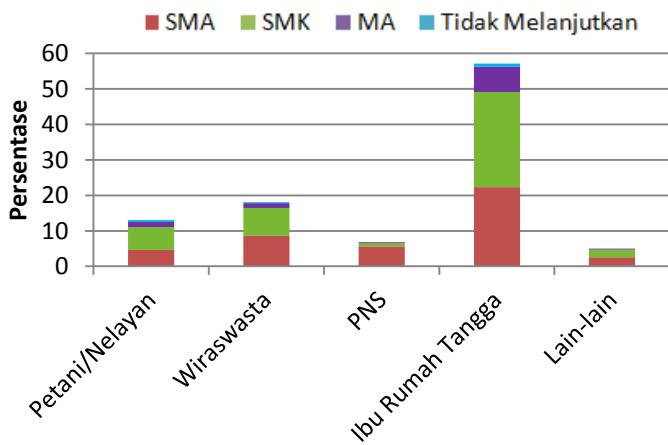


Lampiran 4. Lanjutan

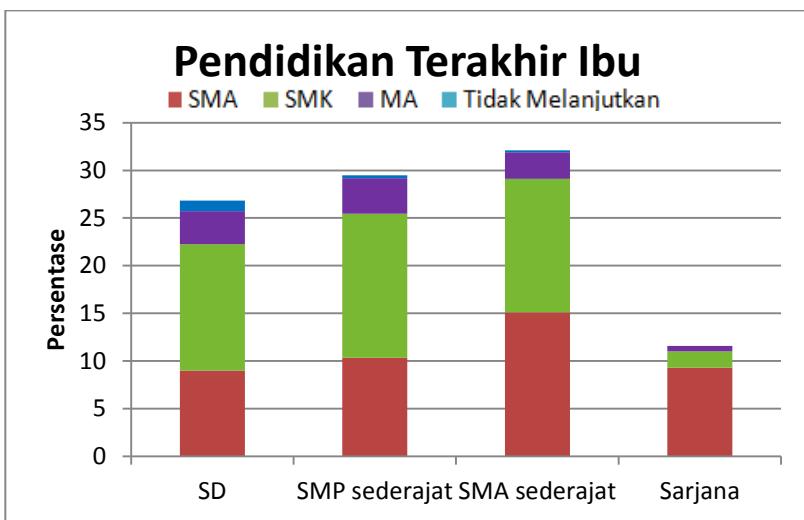
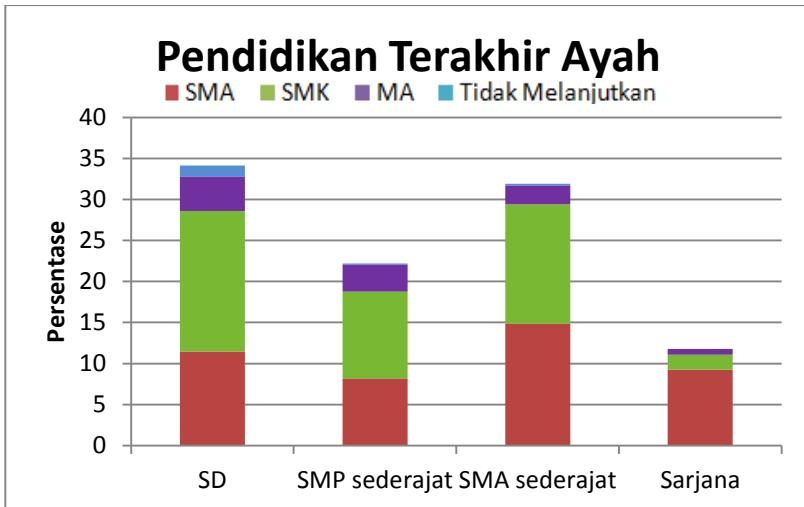
Pekerjaan Ayah



Pekerjaan Ibu

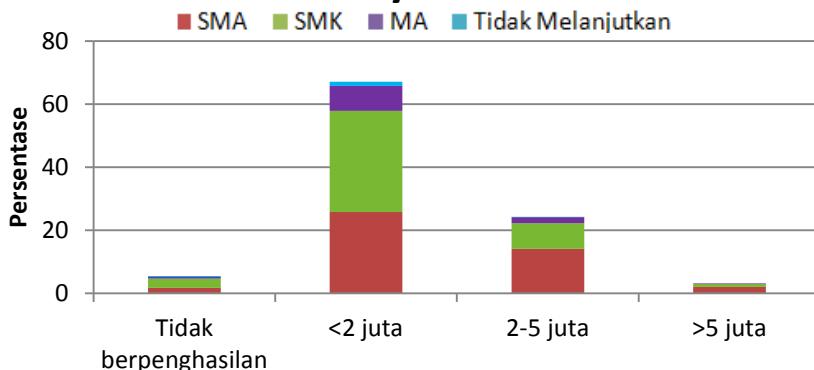


Lampiran 4. Lanjutan

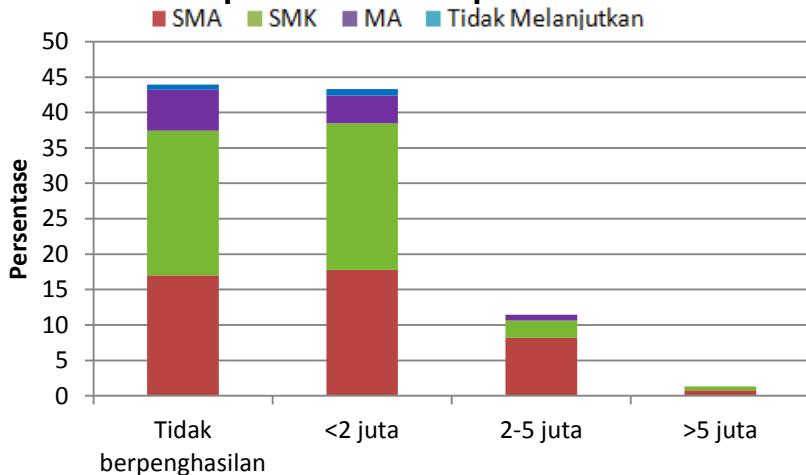


Pendapatan Rata-rata per Bulan

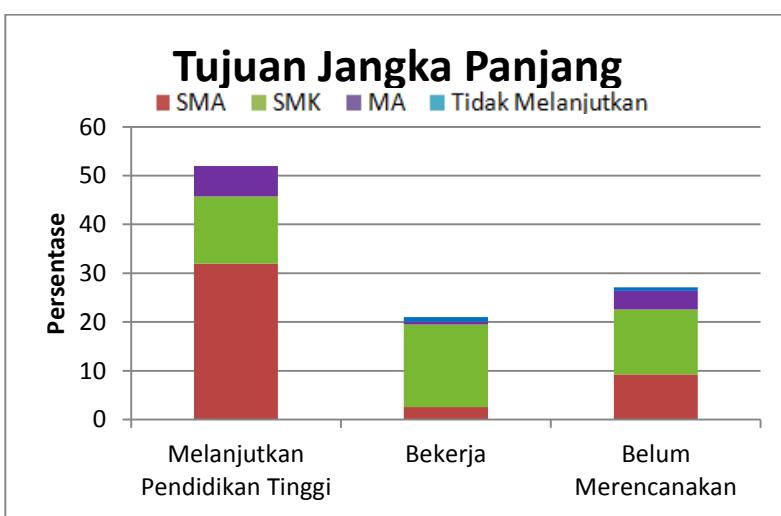
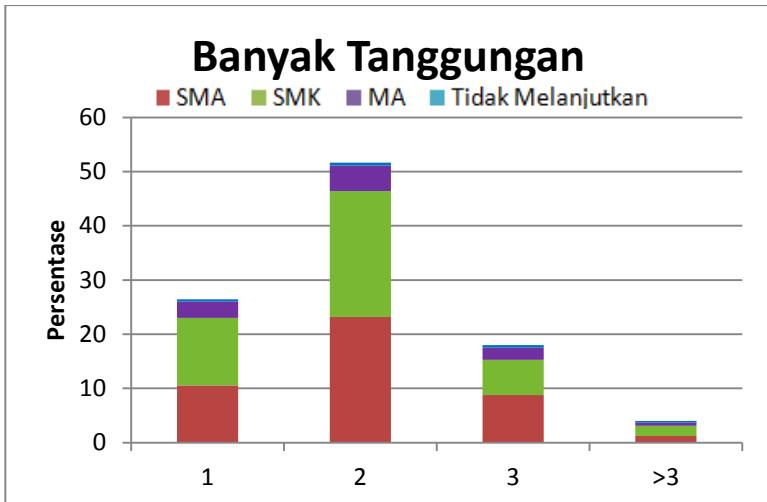
Ayah



Pendapatan Rata-rata per Bulan Ibu

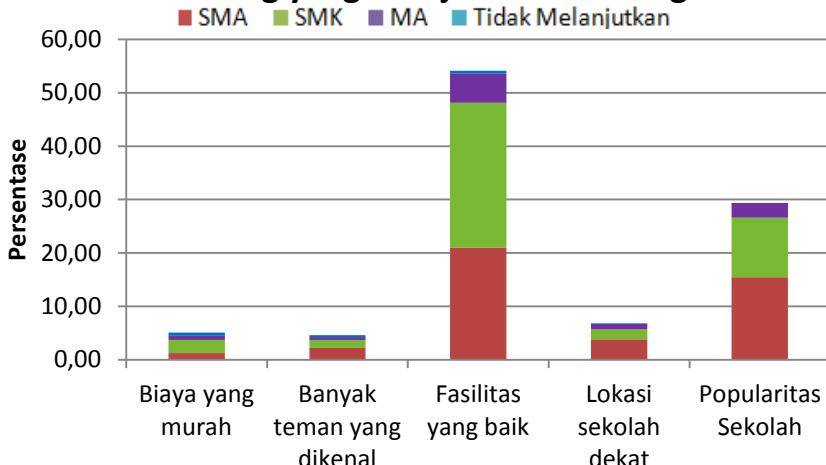


Lampiran 4. Lanjutan

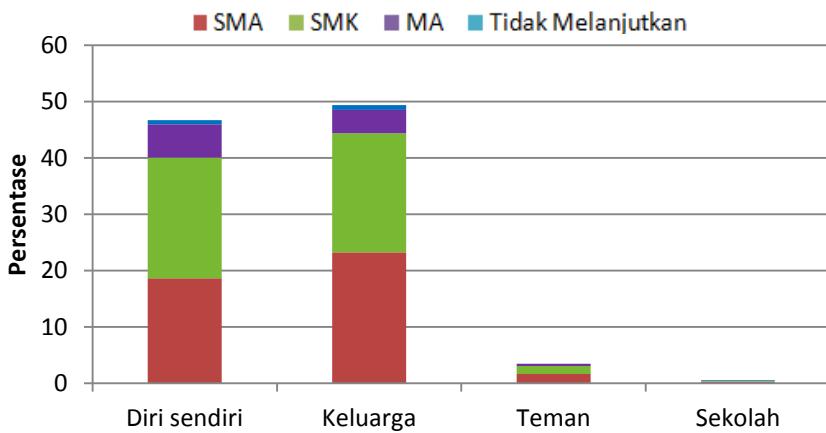


Lampiran 4. Lanjutan

Hal Penting yang Menjadi Pertimbangan



Pihak yang Paling Mempengaruhi Pilihan



Lampiran 5. Surat Keterangan Pengambilan Data di Sekolah

1. SMPN 1 Trenggalek


PEMERINTAH KABUPATEN TRENGGALEK
DINAS PENDIDIKAN, PEMUDA DAN OLAHARGA
SMP NEGERI 1 TRENGGALEK
Jl. Dr. Sutomo No.10 Telp. 0355-791447
TRENGGALEK 66311

SURAT KETERANGAN
Nomor: 070/636/35.03.009.11.201/2017

Yang bertanda tangan dibawah ini:

Nama	:	SUPRIYANTO,S.Pd.
Pangkat/Gol.Ruang	:	Pembina Tk.I, IV/b
N I P	:	19600418 198403 1 005
Jabatan	:	Kepala Sekolah
Unit Kerja	:	SMP Negeri 1 Trenggalek

Menerangkan dengan sebenarnya bahwa:

Nama	:	HAFID WAHYU RAMADHAN
Alamat	:	RT 11 RW 02 Ds. Karangturi Kec. Munjungan, Kab. Trenggalek
Pekerjaan	:	Mahasiswa
Instansi	:	Universitas Brawijaya Malang

Pada tanggal 06 September s.d. 09 September 2017 yang bersangkutan telah melaksanakan penelitian/penyebaran angket pada siswa kelas IX SMP Negeri 1 Trenggalek dalam rangka menyusun skripsi dengan judul:
“Pendugaan Parameter Regresi Logistik Multinomial Menggunakan Metode Least Absolute Shrinkage and Selection Operator (LASSO) Studi Kasus Pilihan Sekolah Siswa SMPN dan MTsN di Kabupaten Trenggalek”.

Demikian Surat Keterangan ini dibuat dan diberikan kepada yang bersangkutan untuk dipergunakan sebagaimana mestinya.

Trenggalek, 09 September 2017
Kepala SMP Negeri 1 Trenggalek,


SUPRIYANTO, S.Pd.
Pembina Tingkat I
NIP. 19600418 198403 1 005

Lampiran 5. Lanjutan

2. SMPN 3 Trenggalek



PEMERINTAH KABUPATEN TRENGGALEK
DINAS PENDIDIKAN, PEMUDA DAN OLAHRAGA
SMP NEGERI 3 TRENGGALEK
Jl. KH Hasyim Asy'ari No. 2 Trenggalek Telp. (0355)791432

SURAT KETERANGAN

Nomor: 421/228/35.03.009. II 203/2017 .

Yang bertandatangan di bawah ini:

Nama : SARDIANI TRI UTOMO, S.Pd., M.Pd
NIP : 196205241985121004
Pangkat/Golongan : Pembina Tk. I / IV b
Unit Kerja : SMP Negeri 3 Trenggalek

Menerangkan dengan sesungguhnya bahwa:

Nama : HAFID WAHYU RAMADHAN
Alamat : RT. 11 RW. 02 Ds. Karangturi, Kec. Munjungan Kab. Trenggalek
Pekerjaan : Mahasiswa S1
Instansi : Universitas Brawijaya Malang
Program Studi : Statistika

Pada tanggal 5 September 2017 yang bersangkutan telah melaksanakan penelitian/penyebaran angket untuk penyusunan skripsi pada siswa kelas IX SMP Negeri 3 Trenggalek.

Demikian Surat Keterangan ini dibuat dan diberikan kepada yang bersangkutan untuk digunakan sebagaimana mestinya.

Trenggalek, 05 September 2017
Kepala SMP Negeri 3 Trenggalek



Lampiran 5. Lanjutan

3. SMPN 5 Trenggalek


PEMERINTAH KABUPATEN TRENGGALEK
DINAS PENDIDIKAN, PEMDA DAN OLAHARGA
SMP NEGERI 5 TRENGGALEK
Alamat : Jl.R.A. Kartini No. 98 Telp. 0355 791388 Trenggalek

SURAT KETERANGAN

Nomor : 422/ 454 /35.03 009.11.205/2017

Yang bertanda tangan di bawah ini :

Nama	:	ANWAR WIYONO, S.Pd
N I P	:	195910131980121003
Jabatan	:	Kepala Sekolah
Unit Kerja	:	SMP Negeri 5 Trenggalek

Menerangkan dengan sebenarnya bahwa :

Nama	:	HAFID WAHYU RAMADHAN
Alamat	:	RT./RW.011 / 022 Ds. Karangturi Kec. Munjungan, Kab. Trenggalek
Asal Intansi	:	Universitas Brawijaya Malang

Telah melaksanakan penelitian guna penyusunan skripsi di SMP Negeri 5 Trenggalek pada hari sabtu tanggal, 9 September 2017

Demikian Surat Keterangan ini dibuat, untuk dapat dipergunakan sebagaimana mestinya.

PEMERINTAH KABUPATEN Trenggalek, 9 September 2017
Kepala Sekolah
SMPN 5 TRENGGALEK
DISDIKPR
ANWAR WIYONO, S.Pd.
NIP. 195910131980121003

Lampiran 5. Lanjutan

4. MTsN 1 Trenggalek

**KEMENTERIAN AGAMA REPUBLIK INDONESIA**
MADRASAH TSANAWIYAH NEGERI
(MTsN) MODEL TRENGGALEK
Alamat : Barat TMP Karangsoko Trenggalek 66314 ☎ (0355) 791562
e-mail: mtsn trenggalek@yahoo.co.id

SURAT KETERANGAN
Nomor : 152/Mts.13.03.01/09/2017

Yang bertanda tangan di bawah ini :

Nama	: Drs. H. AGUNG WIYOTO, M.M.Pd.
NIP	: 196309271988031010
Jabatan	: Kepala Madrasah
Unit Kerja	: MTsN 1 (Model) Trenggalek

Menerangkan dengan sebenarnya bahwa :

Nama	: HAFID WAHYU RAMADHAN
Alamat	: RT 011, RW 022, Desa Karangturi, Kecamatan Munjungan, Kabupaten Trenggalek
Asal Instansi	: Universitas Brawijaya Malang

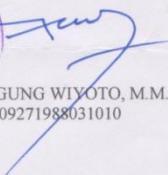
Telah melaksanakan penelitian guna penyusunan skripsi di MTsN 1 (Model) Trenggalek pada hari Sabtu, 09 September 2017.

Demikian surat keterangan ini dibuat dan diberikan kepada yang bersangkutan untuk dipergunakan sebagaimana mestinya.

Trenggalek, 09 September 2017

Kepala MTsN 1 (Model) Trenggalek,

Drs. H. AGUNG WIYOTO, M.M.Pd.
NIP. 196309271988031010



Lampiran 5. Lanjutan

5. SMPN 1 Kampak



PEMERINTAH KABUPATEN TRENGGALEK
DINAS PENDIDIKAN PEMUDA DAN OLAHRAGA
SMP NEGERI 1 KAMPAK
Jln. Anggrek No. 1 Kampak Trenggalek Telp. (0355) 631020

SURAT KETERANGAN

Nomor : 422/219/35.03.009.07.201/2017

Yang Bertanda Tangan dibawah ini :

Nama	: PUJI HARTONO, S.pd, M.pd
NIP	: 19601002 198202 1 001
Jabatan	: Kepala Sekolah
Unit Kerja	: SMP Negeri 1 Kampak

Menerangkan dengan sebenarnya bahwa :

Nama	: HAFID WAHYU RAMADHAN
Alamat	: Rt / Rw. 011 / 022 Ds. Karangturi Kec. Munjungan Kab.Trenggalek
Asal Instansi	: Universitas Brawijaya Malang

Telah melaksanakan penelitian guna penyusunan skripsi di SMP Negeri 1 Kampak pada hari kamis tgl, 07 September 2017

Demikian surat keterangan ini dibuat dan diberikan kepada yang bersangkutan untuk dipergunakan sebagai mana mestinya.

Kampak, 07 September 2017



Lampiran 5. Lanjutan

6. SMPN 2 Kampak


PEMERINTAH KABUPATEN TRENGGALEK
DINAS PENDIDIKAN, PEMUDA DAN OLAHRAGA
SMP NEGERI 2 KAMPAK
RT 13 RW 04 Desa Karangrejo, Kecamatan Kampak, Kabupaten Trenggalek Telp. (0355) 631511

SURAT KETERANGAN
Nomor : 070/113 /35.03.009.07.202/2017

Dasar : Surat Dinas Pendidikan, Pemuda dan Olahraga Kabupaten Trenggalek
Nomor :070/6545/35.03.009/2017 tanggal 4 September 2017 Perihal :
tentang Ijin Penelitian

Yang bertanda tangan di bawah ini :

Nama : **SUYONO, S.Pd. M.Pd.**
NIP : 19611231 198303 1 140
Jabatan : Kepala Sekolah
Unit Kerja : SMP Negeri 2 Kampak

menerangkan dengan sebenarnya bahwa :

Nama : **HAFID WAHYU RAMADHAN**
Alamat : RT/RW 011/002 Ds. Karangturi
Kec. Munjungan, Kab. Trenggalek
Asal Instansi : Universitas Brawijaya Malang
telah melaksanakan penelitian guna penyusunan skripsi di SMP Negeri 2 Kampak pada hari
Kamis tanggal 7 September 2017.

Demikian surat keterangan ini dibuat dan diberikan kepada yang bersangkutan untuk
dipergunakan sebagaimana mestinya.

PEMERINTAH KABUPATEN TRENGGALEK, 7 September 2017
SMPN 2 KAMPAK Kepala SMP Negeri 2 Kampak
* DISDIKOGRA

SUYONO, S.Pd. M.Pd.
NIP. 19611231 198303 1 140

Lampiran 5. Lanjutan

7. SMPN 3 Kampak


PEMERINTAH KABUPATEN TRENGGALEK
DINAS PENDIDIKAN, PEMUDA DAN OLAHRAGA
SMP NEGERI 3 KAMPAK
Jl. Bukit Permai No 1 Timahan Kampak Trenggalek Kode Pos 66373

SURAT KETERANGAN
Nomor : 800/ 081 /35.03.009.07.203/2017

Yang bertanda Tangan dibawah ini :

Nama : **TRISIA ARINI, S.Pd**
Nip. : 196701031989032005
Pangkat/Gol : Pembina Tk.I / IV.b
Jabatan : Kepala Sekolah
Unit Kerja : SMP Negeri 3 Kampak

Menerangkan dengan sebenarnya bahwa :

Nama : **HAFID WAHYU RAMADHAN**
Alamat : Rt/Rw. 011 / 022 Ds Karangturi
Kec. Munjungan Kab. Trenggalek
Asan Instansi : Universitas Brawijaya Malang

Telah Melaksanakan penelitian guna penyusunan skripsi di SMP Negeri 3 Kampak pada hari Jum'at Tanggal, 8 September 2017

Demikian Surat Keterangan ini dibuat dan diberikan kepada yang bersangkutan untuk di pergunakan sebagaimana mestinya.

Kampak 8 September 2017
Kepala Sekolah

DISDIKPA
TRISIA ARINI, S.Pd.
NIP. 196701031989032005

Lampiran 5. Lanjutan

8. MTsN 2 Trenggalek

	<p>KEMENTERIAN AGAMA REPUBLIK INDONESIA KANTOR KEMENTERIAN AGAMA KABUPATEN TRENGGALEK MADRASAH TSANAWIYAH NEGERI KAMPAK Jl. Raya Sugihan Kampak Trenggalek KP 66373 Telp. (0355) 631045 Website : www.mtsnkampak.sch.id email : mtsn.kampak@yahoo.co.id</p>
<p style="text-align: center;">SURAT KETERANGAN Nomor : 362 /MTs.13.03.02/9/2017</p>	
<p>Kepala MTsN Kampak Kabupaten Trenggalek dengan ini menyatakan bahwa :</p>	
<p>Nama : HAFID WAHYU RAMADHAN Program Studi : Statistika NIM : 135090500111008</p>	
<p>Yang bersangkutan benar-benar telah selesai melakukan Penelitian dengan judul Skripsi "Pendugaan Parameter Regresi Logistik Multinomial Menggunakan Metode Least Absolute Shrinkage and Selection Operator (LASSO) Studi Kasus Pilihan Sekolah Siswa SMPN dan MTsN di Kabupaten Trenggalek", pada tanggal 12 September 2017.</p>	
<p>Demikian keterangan ini kami buat untuk dapatnya digunakan sebagaimana mestinya</p>	
<p style="text-align: right;">Trenggalek, 12 September 2017 Kepala,  NASIB SUBANDI</p>	

Lampiran 5. Lanjutan

9. SMPN 1 Munjungan

**PEMERINTAH KABUPATEN TRENGGALEK
DINAS PENDIDIKAN, PEMUDA DAN OLAHARGA
SMP NEGERI 1 MUNJUNGAN**
Jalan. RA. Kartini No. 3 Ds. Masaran Kec. Munjungan Telp (0355) 691053 Kode Pos. 66365

SURAT KETERANGAN
Nomor : 070/374/35.03.009.02.201//2017

Yang bertanda tangan dibawah ini:

Nama	: MOKHAMAD AMIR MAHMUD, M.Pd.
NIP	: 19671221 199703 1 005
Jabatan	: Kepala Sekolah
Unit Kerja	: SMP Negeri 1 Munjungan

Dengan ini menerangkan bahwa :

Nama	: HAFID WAHYU RAMADHAN
Status	: Mahasiswa S1 Uneversitas Brawijaya Malang
NIM	: 135090500111008
Fakultas/Prodi	: MIPA/Statistika

Telah benar-benar melaksanakan penelitian di SMP Negeri 1 Munjungan Kecamatan Munjungan Kabupaten Trenggalek mulai tanggal 11 s.d 12 September 2017 untuk keperluan penyusunan skripsi yang berjudul :

Pendugaan Parameter Regresi Logistik Multinomial Menggunakan Metode Least Absolute Shrinkage and Selection Operator (LASSO) Studi Kasus Pilihan Sekolah Siswa SMPN dan MTsN di Kabupaten Trenggalek.

Demikian surat keterangan ini dibuat dengan sesungguhnya untuk dapat dipergunakan sebagaimana mestinya.

Munjungan, 12 September 2017
Kepada Sekolah,


MOKHAMAD AMIR MAHMUD, M.Pd
Penolong TK I
NIP. 19671221 199703 1 005

Lampiran 5. Lanjutan

10. SMPN 2 Munjungan



PEMERINTAH KABUPATEN TRENGGALEK
DINAS PENDIDIKAN, PEMUDA DAN OLAHRAGA
SMP NEGERI 2 MUNJUNGAN
Desa Ngulungkulon Kec. Munjungan Kab. Trenggalek (66365)

SURAT KETERANGAN

No : 421.72/291/35.03.009.02.202/2017

Yang bertanda tangan di bawah ini, Kepala SMP Negeri 2 Munjungan menerangkan bahwa :

Nama : Hafid Wahyu Ramadhan
NPM : 135090500111008
Perguruan Tinggi : Universitas Brawijaya Malang
FAK-Prodi : MIPA-Statistika
Judul Skripsi : Pendugaan Paramater Regresi Logistik Multinomial
Menggunakan Metode Least Absolute Shrinkage and Selection Operator (LASSO) Studi Kasus Pilihan Sekolah Siswa SMPN
dan MTsN di Kabupaten Trenggalek

Mahasiswa tersebut di atas, telah melaksanakan Survey/Penyebaran Angket untuk penyelesaian skripsi di SMP Negeri 2 Munjungan pada tanggal 13 September 2017.

Demikian Surat Keterangan ini dibuat, untuk dipergunakan sebagaimana mestinya.

Trenggalek, 13 September 2017
Kepala Sekolah
SMPN 2 MUNJUNGAN
DISDIKPO

Drs. AGUS SUYITNO, M.Pd
Pembina
NIP. 19670818 199512 1 004

Lampiran 5. Lanjutan

11. SMPN 3 Munjungan

 <p>PEMERINTAH KABUPATEN TRENGGALEK DINAS PENDIDIKAN, PEMUDA DAN OLAHARGA SMP NEGERI 3 MUNJUNGAN Jl. Raya Munjungan-Watulima Ds. Bendoro Kec. Munjungan Email:smpn3munjungan@gmail.com TRENGGALEK 66365</p>															
<p>SURAT KETERANGAN Nomor : 421/ /35.03.009.02.203/2017</p> <p>Yang bertanda tangan di bawah ini Kepala SMP Negeri 3 Munjungan Kabupaten Trenggalek, menerangkan dengan sesungguhnya bahwa :</p> <table><tr><td>Nama</td><td>:</td><td>HAFID WAHYU RAMADHAN</td></tr><tr><td>Tempat/tanggal lahir</td><td>:</td><td>Trenggalek, 08 Februari 1995</td></tr><tr><td>NIM/NPM</td><td>:</td><td>135090500111008</td></tr><tr><td>Fakultas/Prodi</td><td>:</td><td>MIPA/Statistika</td></tr><tr><td>Instansi</td><td>:</td><td>Universitas Brawijaya Malang</td></tr></table> <p>Telah melakukan penelitian di SMP Negeri 3 Munjungan pada tanggal 4 s/d 30 September 2017, sebagai data penyusun skripsi dengan judul “Pendugaan Parameter Regresi Logistik Multinomial Menggunakan Metode Least Absolute Shrinkage and Selection Operator (LASSO) Studi Kasus Pilihan Sekolah Siswa SMPN dan MTsN di Kabupaten Trenggalek”.</p> <p>Demikian surat keterangan ini kami buat untuk dapat dipergunakan sebagaimana mestinya.</p> <p style="text-align: right;">Trenggalek, 2 Oktober 2017 Kepala Sekolah,  Drs. MUHADI NIP. 19680910 199512 1 002</p>	Nama	:	HAFID WAHYU RAMADHAN	Tempat/tanggal lahir	:	Trenggalek, 08 Februari 1995	NIM/NPM	:	135090500111008	Fakultas/Prodi	:	MIPA/Statistika	Instansi	:	Universitas Brawijaya Malang
Nama	:	HAFID WAHYU RAMADHAN													
Tempat/tanggal lahir	:	Trenggalek, 08 Februari 1995													
NIM/NPM	:	135090500111008													
Fakultas/Prodi	:	MIPA/Statistika													
Instansi	:	Universitas Brawijaya Malang													

Lampiran 5. Lanjutan

12. MTsN 3 Trenggalek



KEMENTERIAN AGAMA REPUBLIK INDONESIA
KANTOR KEMENTERIAN AGAMA KABUPATEN TRENGGALEK
MADRASAH TSANAWIYAH NEGERI 3 TRENGGALEK
Desa Munjungan Kecamatan Munjungan Kabupaten Trenggalek
Telp./Fax 0355-691084 KP. 66365 Email: mtsn.munjungan@yahoo.co.id Website : www.mtsnmunjungan.sch.id

SURAT KETERANGAN

Nomor : B- 217/Mts.13.03.03/PP.00.2/09/2017

Yang bertanda tangan di bawah ini :

Nama : **MUFIDZ, S.Ag, MSI**
N I P : 19681231 199303 1 023
Pangkat/Golongan : Pembina, IV/a
Jabatan : Kepala M Ts Negeri 3 Trenggalek

Dengan ini menerangkan dengan sebenarnya bahwa,

N a m a : **HAFID WAHYU RAMADAN**
Program Studi : STATISTIKA
N I M : 135090500111008

Yang bersangkutan benar-benar telah selesai melakukan Penelitian dengan judul Skripsi "Pendugaan Parameter Regresi Logistik Multinomial Menggunakan Metode *Least Absolute Shrinkage and Selection Operator* (LASSO) Studi Kasus Pilihan Sekolah SMPN dan MTsN di Kabupaten Trenggalek", pada tanggal 13 September 2017.

Demikian surat keterangan ini kami buat untuk dapat dipergunakan sebagaimana mestinya.

Trenggalek, 13 September 2017
Kepala

MUFIDZ, S.Ag, MSI
NIP. 196812311993031023

Lampiran 6. Nilai VIF

1. Nilai VIF Masing-masing Prediktor

Regresi Linier X1 ~ X2 X3 X4 X5 X6 X7 X8 X9 X10 X11 X12

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.378 ^a	.143	.137	3.99801

a. Predictors: (Constant), Pengaruh, Pendidikan_Ayah, Jenis_Sekolah, Pekerjaan_Ibu, Banyak_Tanggungan, Hal_Penting, Tujuan, Pendapatan_Ibu, Pendapatan_Ayah, Pekerjaan_Ayah, Pendidikan_Ibu

$$VIF = \frac{1}{1 - R^2} = \frac{1}{1 - 0,143} = 1,17$$

Regresi Logistik X2 ~ X1 X3 X4 X5 X6 X7 X8 X9 X10 X11 X12

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1808.639 ^a	.117	.161

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

$$VIF = \frac{1}{1 - R^2} = \frac{1}{1 - 0,117} = 1,13$$

Regresi Logistik X3 ~ X1 X2 X4 X5 X6 X7 X8 X9 X10 X11 X12

Pseudo R-Square

Cox and Snell	.578
Nagelkerke	.632
McFadden	.350

$$VIF = \frac{1}{1 - R^2} = \frac{1}{1 - 0,578} = 2,37$$

Regresi Logistik X4 ~ X1 X2 X3 X5 X6 X7 X8 X9 X10 X11 X12

Pseudo R-Square

Cox and Snell	.680
Nagelkerke	.743
McFadden	.464

$$VIF = \frac{1}{1 - R^2} = \frac{1}{1 - 0,680} = 3,13$$

Regresi Logistik X5 ~ X1 X2 X3 X4 X6 X7 X8 X9 X10 X11 X12

Pseudo R-Square

Cox and Snell	.582
Nagelkerke	.627
McFadden	.331

$$VIF = \frac{1}{1 - R^2} = \frac{1}{1 - 0,582} = 2,39$$

Regresi Logistik X6 ~ X1 X2 X3 X4 X5 X7 X8 X9 X10 X11 X12

Pseudo R-Square

Cox and Snell	.585
Nagelkerke	.629
McFadden	.331

$$VIF = \frac{1}{1 - R^2} = \frac{1}{1 - 0,585} = 2,41$$

Regresi Logistik X7 ~ X1 X2 X3 X4 X5 X6 X8 X9 X10 X11 X12

Pseudo R-Square

Cox and Snell	.310
Nagelkerke	.374
McFadden	.211

$$VIF = \frac{1}{1 - R^2} = \frac{1}{1 - 0,310} = 1,45$$

Regresi Logistik X8 ~ X1 X2 X3 X4 X5 X6 X7 X9 X10 X11 X12

Pseudo R-Square

Cox and Snell	.467
Nagelkerke	.536
McFadden	.306

$$VIF = \frac{1}{1 - R^2} = \frac{1}{1 - 0,467} = 1,88$$

Regresi Logistik X9 ~ X1 X2 X3 X4 X5 X6 X7 X8 X10 X11 X12

Pseudo R-Square

Cox and Snell	.046
Nagelkerke	.052
McFadden	.021

$$VIF = \frac{1}{1 - R^2} = \frac{1}{1 - 0,046} = 1,05$$

Regresi Logistik X10 ~ X1 X2 X3 X4 X5 X6 X7 X8 X9 X11 X12

Pseudo R-Square

Cox and Snell	.224
Nagelkerke	.257
McFadden	.124

$$VIF = \frac{1}{1 - R^2} = \frac{1}{1 - 0,224} = 1,29$$

Regresi Logistik X11 ~ X1 X2 X3 X4 X5 X6 X7 X8 X9 X10 X12

Pseudo R-Square

Pseudo R-Square	
Cox and Snell	.203
Nagelkerke	.225
McFadden	.097

$$VIF = \frac{1}{1 - R^2} = \frac{1}{1 - 0,203} = 1,25$$

Regresi Logistik X12 ~ X1 X2 X3 X4 X5 X6 X7 X8 X9 X10 X11

Pseudo R-Square

Pseudo R-Square	
Cox and Snell	.077
Nagelkerke	.095
McFadden	.048

$$VIF = \frac{1}{1 - R^2} = \frac{1}{1 - 0,077} = 1,08$$

Lampiran 7. Output SPSS 22 untuk Persentase

Case Processing Summary

		N	Marginal Percentage
Pilihan	SMA	672	43.8%
	SMK	677	44.1%
	MA	162	10.5%
	Tidak Melanjutkan	25	1.6%
Jenis_Sekolah	SMP	989	64.4%
	MTs	547	35.6%
Pekerjaan_Ayah	Petani/Nelayan	636	41.4%
	Wiraswasta	539	35.1%
	PNS	146	9.5%
	Lain-Lain	215	14.0%
Pekerjaan_Ibu	Petani/Nelayan	200	13.0%
	Wiraswasta	277	18.0%
	PNS	105	6.8%
	Ibu Rumah Tangga	877	57.1%
	Lain-lain	77	5.0%
Pendidikan_Ayah	SD	524	34.1%
	SMP Sederajat	341	22.2%
	SMA sederajat	490	31.9%
	Perguruan Tinggi	181	11.8%
Pendidikan_Ibu	SD	412	26.8%
	SMP Sederajat	453	29.5%
	SMA sederajat	493	32.1%
	Perguruan Tinggi	178	11.6%
Pendapatan_Ayah	Tidak Berpenghasilan	85	5.5%
	<2 juta	1030	67.1%
	2-5 juta	372	24.2%
	>5 juta	49	3.2%

Lampiran 7. Lanjutan

Pendapatan_Ibu	Tidak Berpenghasilan	675	43.9%
	<2 juta	665	43.3%
	2-5 juta	176	11.5%
	>5 juta	20	1.3%
Banyak_Tanggungan	1 anak	406	26.4%
	2 anak	793	51.6%
	3 anak	276	18.0%
	> 3 anak	61	4.0%
Tujuan	Melanjutkan	798	52.0%
	Perguruan Tinggi		
	Langsung Bekerja	322	21.0%
	Belum	416	27.1%
	Merencanakan		
Hal_Penting	Biaya Murah	78	5.1%
	Banyak teman yang dikenal	71	4.6%
	Fasilitas yang baik	831	54.1%
	Lokasi sekolah dekat	105	6.8%
	Popularitas/Sekolah Favorit	451	29.4%
Pengaruh	Tidak Ada/Diri Sendiri	717	46.7%
	Keluarga	758	49.3%
	Teman	53	3.5%
	Sekolah	8	0.5%
Valid		1536	100.0%
Missing		0	
Total		1536	
Subpopulation		1503 ^a	

a. The dependent variable has only one value observed in 1489 (99.1%) subpopulations.

Lampiran 8. Output SPSS 22 untuk Uji Simultan dan Uji Kesesuaian Model

Output untuk uji parameter secara simultan:

Model Fitting Information

Model	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	3140,812	3156,823	3134,812			
Final	2430,500	2974,867	2226,500	908,312	99	,000

Output untuk uji kesesuaian model:

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	3539.434	4407	1.000
Deviance	2206.856	4407	1.000

Lampiran 9. Output SPSS 22 untuk Pendugaan Parameter, Uji Parsial dan *Odds Ratio*

Parameter Estimates

Pilihan ^a		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
SMA	Intercept	55.465	2333.224	.001	1	.981			
	Nilai_Rata2	.240	.066	13.410	1	.000	1.271	1.118	1.446
	[Jenis_Sekolah=0]	.125	.608	.042	1	.837	1.133	.344	3.728
	[Jenis_Sekolah=1]	0 ^b	.	.	0
	[Pekerjaan_Ayah=0]	-16.575	486.311	.001	1	.973	6.334E-8	.000	c
	[Pekerjaan_Ayah=1]	-15.283	486.311	.001	1	.975	2.305E-7	.000	c
	[Pekerjaan_Ayah=2]	-3.670	781.993	.000	1	.996	.025	.000	c
	[Pekerjaan_Ayah=3]	0 ^b	.	.	0
	[Pekerjaan_Ibu=0]	-13.939	631.710	.000	1	.982	8.843E-7	.000	c
	[Pekerjaan_Ibu=1]	-14.605	631.710	.001	1	.982	4.542E-7	.000	c

Lampiran 9. Lanjutan.

[Pekerjaan_Ibu=2]	-13.836	1138.870	.000	1	.990	9.800E-7	.000	c
[Pekerjaan_Ibu=3]	-13.428	631.710	.000	1	.983	1.473E-6	.000	c
[Pekerjaan_Ibu=4]	0 ^b	.	.	0
[Pendidikan_Ayah=0]	-10.917	586.100	.000	1	.985	1.814E-5	.000	c
[Pendidikan_Ayah=1]	-9.224	586.100	.000	1	.987	9.865E-5	.000	c
[Pendidikan_Ayah=2]	-10.798	586.100	.000	1	.985	2.045E-5	.000	c
[Pendidikan_Ayah=3]	0 ^b	.	.	0
[Pendidikan_Ibu=0]	-9.473	555.962	.000	1	.986	7.687E-5	.000	c
[Pendidikan_Ibu=1]	-9.158	555.962	.000	1	.987	.000	.000	c
[Pendidikan_Ibu=2]	-9.225	555.962	.000	1	.987	9.853E-5	.000	c
[Pendidikan_Ibu=3]	0 ^b	.	.	0
[Pendapatan_Ayah=0]	-10.927	1526.010	.000	1	.994	1.796E-5	.000	c
[Pendapatan_Ayah=1]	-10.056	1526.009	.000	1	.995	4.292E-5	.000	c
[Pendapatan_Ayah=2]	-10.378	1526.010	.000	1	.995	3.112E-5	.000	c
[Pendapatan_Ayah=3]	0 ^b	.	.	0

Lampiran 9. Lanjutan.

[Pendapatan_Ibu=0]	-1.554	2084.111	.000	1	.999	.211	.000	.	^c
[Pendapatan_Ibu=1]	-1.237	2084.111	.000	1	1.000	.290	.000	.	^c
[Pendapatan_Ibu=2]	12.014	2082.623	.000	1	.995	165105.410	.000	.	^c
[Pendapatan_Ibu=3]	0 ^b	.	.	0
[Banyak_Tanggungan=0]	4.024	1.016	15.692	1	.000	55.920	7.637	409.482	
[Banyak_Tanggungan=1]	4.425	1.029	18.487	1	.000	83.535	11.112	627.982	
[Banyak_Tanggungan=2]	3.429	1.040	10.881	1	.001	30.846	4.021	236.603	
[Banyak_Tanggungan=3]	0 ^b	.	.	0
[Tujuan=0]	15.714	271.273	.003	1	.954	6672731.86	8.250	5.397E+237	
[Tujuan=1]	-2.616	.638	16.795	1	.000	.073	.021	.255	
[Tujuan=2]	0 ^b	.	.	0	
[Hal_Penting=0]	-16.965	477.174	.001	1	.972	4.289E-8	.000	.	^c
[Hal_Penting=1]	-15.679	477.174	.001	1	.974	1.551E-7	.000	.	^c
[Hal_Penting=2]	-14.055	477.174	.001	1	.977	7.868E-7	.000	.	^c
[Hal_Penting=3]	-14.336	477.175	.001	1	.976	5.940E-7	.000	.	^c

Lampiran 9. Lanjutan.

	[Hal_Penting=4]	0 ^b	.	.	0
	[Pengaruh=0]	.257	2.474	.011	1	.917	1.293	.010	164.965
	[Pengaruh=1]	-.073	2.443	.001	1	.976	.930	.008	111.606
	[Pengaruh=2]	13.913	752.227	.000	1	.985	1102646.113	.000	.
	[Pengaruh=3]	0 ^b	.	.	0
SMK	Intercept	58.784	2333.224	.001	1	.980	.	.	.
	Nilai_Rata2	.175	.064	7.451	1	.006	1.192	1.051	1.352
	[Jenis_Sekolah=0]	-.166	.599	.076	1	.782	.847	.262	2.740
	[Jenis_Sekolah=1]	0 ^b	.	.	0
	[Pekerjaan_Ayah=0]	-16.607	486.311	.001	1	.973	6.132E-8	.000	.
	[Pekerjaan_Ayah=1]	-15.043	486.311	.001	1	.975	2.929E-7	.000	.
	[Pekerjaan_Ayah=2]	-3.980	781.993	.000	1	.996	.019	.000	.
	[Pekerjaan_Ayah=3]	0 ^b	.	.	0
	[Pekerjaan_Ibu=0]	-13.871	631.710	.000	1	.982	9.461E-7	.000	.
	[Pekerjaan_Ibu=1]	-14.730	631.710	.001	1	.981	4.005E-7	.000	.
	[Pekerjaan_Ibu=2]	-13.887	1138.870	.000	1	.990	9.313E-7	.000	.
	[Pekerjaan_Ibu=3]	-13.493	631.710	.000	1	.983	1.381E-6	.000	.
	[Pekerjaan_Ibu=4]	0 ^b	.	.	0

Lampiran 9. Lanjutan.

[Pendidikan_Ayah=0]	-10.267	586.100	.000	1	.986	3.477E-5	.000	.
[Pendidikan_Ayah=1]	-8.573	586.100	.000	1	.988	.000	.000	c
[Pendidikan_Ayah=2]	-10.063	586.100	.000	1	.986	4.264E-5	.000	c
[Pendidikan_Ayah=3]	0 ^b	.	.	0
[Pendidikan_Ibu=0]	-8.801	555.962	.000	1	.987	.000	.000	c
[Pendidikan_Ibu=1]	-8.327	555.962	.000	1	.988	.000	.000	c
[Pendidikan_Ibu=2]	-8.657	555.962	.000	1	.988	.000	.000	c
[Pendidikan_Ibu=3]	0 ^b	.	.	0
[Pendapatan_Ayah=0]	-10.969	1526.010	.000	1	.994	1.723E-5	.000	c
[Pendapatan_Ayah=1]	-10.359	1526.009	.000	1	.995	3.170E-5	.000	c
[Pendapatan_Ayah=2]	-10.765	1526.010	.000	1	.994	2.113E-5	.000	c
[Pendapatan_Ayah=3]	0 ^b	.	.	0
[Pendapatan_Ibu=0]	-1.188	2084.111	.000	1	1.000	.305	.000	c
[Pendapatan_Ibu=1]	-.862	2084.111	.000	1	1.000	.422	.000	c
[Pendapatan_Ibu=2]	11.869	2082.623	.000	1	.995	142771.976	.000	c
[Pendapatan_Ibu=3]	0 ^b	.	.	0
[Banyak_Tanggungan=0]	3.187	.972	10.753	1	.001	24.216	3.604	162.702
[Banyak_Tanggungan=1]	3.609	.988	13.354	1	.000	36.945	5.331	256.037

Lampiran 9. Lanjutan.

[Banyak_Tanggungan=2]	2.435	.996	5.974	1	.015	11.410	1.620	80.378
[Banyak_Tanggungan=3]	0 ^b	.	.	0
[Tujuan=0]	14.671	271.273	.003	1	.957	2352601.6 57	2.909E-225	1.903E+237
[Tujuan=1]	-.977	.616	2.521	1	.112	.376	.113	1.257
[Tujuan=2]	0 ^b	.	.	0
[Hal_Penting=0]	-16.897	477.174	.001	1	.972	4.588E-8	.000	c
[Hal_Penting=1]	-16.360	477.174	.001	1	.973	7.849E-8	.000	c
[Hal_Penting=2]	-13.711	477.174	.001	1	.977	1.110E-6	.000	c
[Hal_Penting=3]	-15.497	477.175	.001	1	.974	1.861E-7	.000	c
[Hal_Penting=4]	0 ^b	.	.	0
[Pengaruh=0]	2.358	2.538	.863	1	.353	10.571	.073	1528.878
[Pengaruh=1]	1.887	2.508	.566	1	.452	6.599	.048	900.001
[Pengaruh=2]	15.505	752.227	.000	1	.984	5417406.4 30	.000	c
[Pengaruh=3]	0 ^b	.	.	0
MA	Intercept	45.071	2022.915	.000	1	.982		
	Nilai_Rata2	.181	.068	7.159	1	.007	1.198	1.050
	[Jenis_Sekolah=0]	-3.173	.653	23.644	1	.000	.042	.012
								.150

Lampiran 9. Lanjutan.

[Jenis_Sekolah=1]	0 ^b	.	.	0
[Pekerjaan_Ayah=0]	-16.142	486.311	.001	1	.974	9.764E-8	.000	.	c
[Pekerjaan_Ayah=1]	-15.775	486.312	.001	1	.974	1.410E-7	.000	.	c
[Pekerjaan_Ayah=2]	3.670	1289.282	.000	1	.998	39.262	.000	.	c
[Pekerjaan_Ayah=3]	0 ^b	.	.	0
[Pekerjaan_Ibu=0]	14.939	1041.511	.000	1	.989	3074086.700	.000	.	c
[Pekerjaan_Ibu=1]	15.605	1041.511	.000	1	.988	5985200.952	.000	.	c
[Pekerjaan_Ibu=2]	14.836	1877.674	.000	1	.994	2773613.019	.000	.	c
[Pekerjaan_Ibu=3]	14.428	1041.511	.000	1	.989	1845306.613	.000	.	c
[Pekerjaan_Ibu=4]	0 ^b	.	.	0
[Pendidikan_Ayah=0]	11.917	966.308	.000	1	.990	149847.835	.000	.	c
[Pendidikan_Ayah=1]	10.224	966.309	.000	1	.992	27552.677	.000	.	c
[Pendidikan_Ayah=2]	-10.381	586.100	.000	1	.986	3.102E-5	.000	.	c
[Pendidikan_Ayah=3]	0 ^b	.	.	0
[Pendidikan_Ibu=0]	-8.389	555.963	.000	1	.988	.000	.000	.	c
[Pendidikan_Ibu=1]	-8.186	555.962	.000	1	.988	.000	.000	.	c

Lampiran 9. Lanjutan.

[Pendidikan_Ibu=2]	-8.760	555.962	.000	1	.987	.000	.000	c
[Pendidikan_Ibu=3]	0 ^b	.	.	0
[Pendapatan_Ayah=0]	-11.049	1526.01	.000	1	.994	1.591E-5	.000	c
[Pendapatan_Ayah=1]	-10.194	1526.01	.000	1	.995	3.740E-5	.000	c
[Pendapatan_Ayah=2]	-10.618	1526.01	.000	1	.994	2.448E-5	.000	c
[Pendapatan_Ayah=3]	0 ^b	.	.	0
[Pendapatan_Ibu=0]	15.060	595.800	.001	1	.980	3470728.513	.000	c
[Pendapatan_Ibu=1]	15.178	595.800	.001	1	.980	3904019.827	.000	c
[Pendapatan_Ibu=2]	28.697	.000	.	1	.	2902365.059498.030	2902365059498.030	
[Pendapatan_Ibu=3]	0 ^b	.	.	0
[Banyak_Tanggungan=0]	2.940	1.045	7.914	1	.005	18.925	2.440	146.818
[Banyak_Tanggungan=1]	3.158	1.055	8.954	1	.003	23.525	2.973	186.155
[Banyak_Tanggungan=2]	2.572	1.071	5.765	1	.016	13.093	1.604	106.867

Lampiran 9. Lanjutan.

[Banyak_Tanggungan=3]	0 ^b	.	.	0
[Tujuan=0]	15.594	271.273	.003	1	.954	5920861 .923	7.320E- -225	4.789E+237	
[Tujuan=1]	-2.562	.718	12.723	1	.000	.077	.019	.315	
[Tujuan=2]	0 ^b	.	.	0	
[Hal_Penting=0]	-16.531	477.174	.001	1	.972	6.618E- 8	.000	.	^c
[Hal_Penting=1]	-16.346	477.174	.001	1	.973	7.960E- 8	.000	.	^c
[Hal_Penting=2]	-14.142	477.174	.001	1	.976	7.216E- 7	.000	.	^c
[Hal_Penting=3]	-14.561	477.175	.001	1	.976	4.747E- 7	.000	.	^c
[Hal_Penting=4]	0 ^b	.	.	0	
[Pengaruh=0]	-1.931	2.608	.548	1	.459	.145	.001	24.055	
[Pengaruh=1]	-2.501	2.577	.942	1	.332	.082	.001	12.814	
[Pengaruh=2]	11.828	752.228	.000	1	.987	137025. 460	.000	.	^c
[Pengaruh=3]	0 ^b	.	.	0	

a. The reference category is: SMA.

b. This parameter is set to zero because it is redundant.

c. Floating point overflow occurred while computing this statistic. Its value is therefore set to system missing.

Lampiran 10. *Syntax R* untuk Pendugaan Regresi Logistik Multinomial menggunakan LASSO.

```
#Input Data
data=read.csv(file.choose(),header=T)

#Menampilkan Kategori
data$Y=factor(data$Y,
               levels=c(3,0,1,2),
               labels=c("Tidak
               Melanjutkan","SMA","SMK","MA"))
data$x2=factor(data$x2,
                levels=c(1,0),
                labels=c("MTs","SMP"))
data$x3=factor(data$x3,
                levels=c(3,0,1,2),
                labels=c("Lain-
                lain","Petani/Nelayan","wiraswasta","PNS"))
data$x4=factor(data$x4,
                levels=c(4,0,1,2,3),
                labels=c("Lain-
                lain","Petani/Nelayan","wiraswasta","PNS","I
                bu Rumah Tangga"))
data$x5=factor(data$x5,
                levels=c(3,0,1,2),
                labels=c("Perguruan Tinggi","SD","SMP
                Sederajat","SMA Sederajat"))
data$x6=factor(data$x6,
                levels=c(3,0,1,2),
                labels=c("Perguruan Tinggi","SD","SMP
                Sederajat","SMA Sederajat"))
data$x7=factor(data$x7,
                levels=c(3,0,1,2),
                labels=c(">5 juta","Tidak
                Berpenghasilan",<2 juta","2-5 juta"))
data$x8=factor(data$x8,
                levels=c(3,0,1,2),
                labels=c(">5 juta","Tidak
                Berpenghasilan",<2 juta","2-5 juta"))
data$x9=factor(data$x9,
                levels=c(3,0,1,2),
                labels=c(">3 anak","1 anak","2 anak","3
                anak"))
data$x10=factor(data$x10,
                levels=c(2,0,1),
                labels=c("Belum
                merencanakan","Melanjutkan Pendidikan
                Tinggi","Langsung bekerja"))
```

Lampiran 10. Lanjutan

```
data$x11=factor(data$x11,
                  levels=c(4,0,1,2,3),
                  labels=c("Popularitas Sekolah/Sekolah
                           Favorit","Biaya yang Murah","Banyak teman
                           yang dikenal","Fasilitas yang baik","Lokasi
                           dekat dengan rumah"))
data$x12=factor(data$x12,
                  levels=c(3,0,1,2),
                  labels=c("Sekolah","Tidak Ada/Diri
                           Sendiri","Keluarga","Teman"))

Y=data[,13]
X1=data[,1]
X2=data[,2]
X3=data[,3]
X4=data[,4]
X5=data[,5]
X6=data[,6]
X7=data[,7]
X8=data[,8]
X9=data[,9]
X10=data[,10]
X11=data[,11]
X12=data[,12]

#Membuat Matriks X
Xfactors=model.matrix(Y~X2+X3+X4+X5+X6+X7+X8+X9+X10+X11
                      +X12)[,-1]
X=cbind(X1,Xfactors)
Ydummy=model.matrix(~Y)[,-1]

#Pendugaan parameter menggunakan Lasso
library(glmnet)
cvfit = cv.glmnet(x,Ydummy,
                   alpha=1,family="multinomial")
cvfit
fit = glmnet(x,Ydummy,
              family="multinomial",lambda=0.02588615)
coef(fit)

#Mencari nilai AIC
t11=fit$nulldev - deviance(fit) #nilai -2log likelihood
k=102
n=1536
AIC=2*k+t11
```

Lampiran 11. Output R Pendugaan Parameter Regresi Logistik Multinomial menggunakan LASSO.

```
> cvfit
$lambda.min
[1] 0.004027047

$lambda.1se
[1] 0.02588615

> fit =
glmnet(x,ydummy,family="multinomial",lambda=0.02588615)
> coef(fit)
$YsMA
34 x 1 sparse Matrix of class "dgCMatrix"
           s0
x1          -2.05332615
x2SMP        0.03787966
x3Petani/Nelayan .
x3wiraswasta .
x3PNS         .
x4Petani/Nelayan .
x4wiraswasta .
x4PNS         0.47329307
x4Ibu Rumah Tangga .
x5SD          .
x5SMP Sederajat .
x5SMA Sederajat .
x6SD          .
x6SMP Sederajat -0.09829057
x6SMA Sederajat .
x7Tidak Berpenghasilan .
x7<2 juta    .
x72-5 juta   0.02526744
x8Tidak Berpenghasilan .
x8<2 juta    .
x82-5 juta   0.30024937
x91 anak     .
x92 anak     .
x93 anak     .
x10Melanjutkan Pendidikan Tinggi 0.32800471
x10Langsung bekerja .
x11Biaya yang Murah .
x11Banyak teman yang dikenal .
x11Fasilitas yang baik .
```

Lampiran 11. Lanjutan.

X11Lokasi dekat dengan rumah	.
X12Tidak Ada/Diri Sendiri	.
X12Keluarga	.
X12Teman	.
 \$YSMK	
34 x 1 sparse Matrix of class "dgCMatrix"	
	s0
	1.3078232
X1	.
X2SMP	.
X3Petani/Nelayan	.
X3Wiraswasta	0.0234915
X3PNS	-0.2367372
X4Petani/Nelayan	.
X4Wiraswasta	.
X4PNS	.
X4Ibu Rumah Tangga	.
X5SD	.
X5SMP Sederajat	.
X5SMA Sederajat	.
X6SD	.
X6SMP Sederajat	.
X6SMA Sederajat	.
X7Tidak Berpenghasilan	.
X7<2 juta	.
X72-5 juta	.
X8Tidak Berpenghasilan	.
X8<2 juta	.
X82-5 juta	.
X91 anak	.
X92 anak	.
X93 anak	.
X10Melanjutkan Pendidikan Tinggi	-0.5602946
X10Langsung bekerja	1.2557468
X11Biaya yang Murah	.
X11Banyak teman yang dikenal	.
X11Fasilitas yang baik	0.2093564
X11Lokasi dekat dengan rumah	-0.4728261
X12Tidak Ada/Diri Sendiri	.
X12Keluarga	.
X12Teman	.

Lampiran 11. Lanjutan.

```
$YMA
34 x 1 sparse Matrix of class "dgCMatrix"
           s0
X1          .
X2SMP      -2.0278484
X3Petani/Nelayan 0.2351000
X3Wiraswasta   .
X3PNS        .
X4Petani/Nelayan  .
X4Wiraswasta   .
X4PNS        .
X4Ibu Rumah Tangga  .
X5SD         .
X5SMP Sederajat  .
X5SMA Sederajat  .
X6SD         .
X6SMP Sederajat  .
X6SMA Sederajat  .
X7Tidak Berpenghasilan  .
X7<2 juta    .
X72-5 juta   .
X8Tidak Berpenghasilan  .
x8<2 juta    .
x82-5 juta   .
x91 anak     .
x92 anak     .
x93 anak     .
x10Melanjutkan Pendidikan Tinggi  .
x10Langsung bekerja  .
x11Biaya yang Murah  .
x11Banyak teman yang dikenal  .
x11Fasilitas yang baik  .
x11Lokasi dekat dengan rumah  .
x12Tidak Ada/Diri Sendiri  .
x12Keluarga    .
x12Teman      .

> AIC
[1] 837.1146
```


Lampiran 12. Perbandingan Koefisien dan Signifikansi Parameter antara Metode MLE dan LASSO

1. Logit 1(SMA)

Penduga Parameter	MLE		LASSO	
	Koefisien	Signifikansi	Koefisien	Signifikansi
β_{10}	55,465	Tidak	-2,053	Signifikan
β_{11}	0,24	Signifikan	0,038	Signifikan
$\beta_{12(0)}$	0,125	Tidak	0,015	Signifikan
$\beta_{13(0)}$	-16,575	Tidak	0	Tidak
$\beta_{13(1)}$	-15,283	Tidak	0	Tidak
$\beta_{13(2)}$	-3,67	Tidak	0	Tidak
$\beta_{14(0)}$	-13,939	Tidak	0	Tidak
$\beta_{14(1)}$	-14,605	Tidak	0	Tidak
$\beta_{14(2)}$	-13,836	Tidak	0,473	Signifikan
$\beta_{14(3)}$	-13,428	Tidak	0	Tidak
$\beta_{15(0)}$	-10,917	Tidak	0	Tidak
$\beta_{15(1)}$	-9,224	Tidak	0	Tidak
$\beta_{15(2)}$	-10,798	Tidak	0	Tidak
$\beta_{16(0)}$	-9,473	Tidak	0	Tidak
$\beta_{16(1)}$	-9,158	Tidak	-0,098	Signifikan
$\beta_{16(2)}$	-9,225	Tidak	0	Tidak
$\beta_{17(0)}$	-10,927	Tidak	0	Tidak
$\beta_{17(1)}$	-10,056	Tidak	0	Tidak
$\beta_{17(2)}$	-10,378	Tidak	0,025	Signifikan
$\beta_{18(0)}$	-1,554	Tidak	0	Tidak
$\beta_{18(1)}$	-1,237	Tidak	0	Tidak
$\beta_{18(2)}$	12,014	Tidak	0,3	Signifikan
$\beta_{19(0)}$	4,024	Signifikan	0	Tidak
$\beta_{19(1)}$	4,425	Signifikan	0	Tidak
$\beta_{19(2)}$	3,429	Signifikan	0	Tidak
$\beta_{110(0)}$	15,714	Tidak	0,328	Signifikan
$\beta_{110(1)}$	-2,616	Signifikan	0	Tidak

Lampiran 12. Lanjutan

Penduga Parameter	MLE		LASSO	
	Koefisien	Signifikansi	Koefisien	Signifikansi
$\beta_{111(0)}$	-16,965	Tidak	0	Tidak
$\beta_{111(1)}$	-15,679	Tidak	0	Tidak
$\beta_{111(2)}$	-14,055	Tidak	0	Tidak
$\beta_{111(3)}$	-14,336	Tidak	0	Tidak
$\beta_{112(0)}$	0,257	Tidak	0	Tidak
$\beta_{112(1)}$	-0,073	Tidak	0	Tidak
$\beta_{112(2)}$	13,913	Tidak	0	Tidak

2. Logit 2(SMK)

Penduga Parameter	MLE		Koefisien LASSO	
	Koefisien	Signifikansi	Koefisien	Signifikansi
β_{10}	58,784	Tidak	1,308	Signifikan
β_{11}	0,175	Signifikan	0	Tidak
$\beta_{12(0)}$	-0,166	Tidak	0	Tidak
$\beta_{13(0)}$	-16,607	Tidak	0	Tidak
$\beta_{13(1)}$	-15,043	Tidak	0,023	Signifikan
$\beta_{13(2)}$	-3,98	Tidak	-0,237	Signifikan
$\beta_{14(0)}$	-13,871	Tidak	0	Tidak
$\beta_{14(1)}$	-14,73	Tidak	0	Tidak
$\beta_{14(2)}$	-13,887	Tidak	0	Tidak
$\beta_{14(3)}$	-13,493	Tidak	0	Tidak
$\beta_{15(0)}$	-10,267	Tidak	0	Tidak
$\beta_{15(1)}$	-8,573	Tidak	0	Tidak
$\beta_{15(2)}$	-10,063	Tidak	0	Tidak
$\beta_{16(0)}$	-8,801	Tidak	0	Tidak
$\beta_{16(1)}$	-8,327	Tidak	0	Tidak
$\beta_{16(2)}$	-8,657	Tidak	0	Tidak
$\beta_{17(0)}$	-10,969	Tidak	0	Tidak
$\beta_{17(1)}$	-10,359	Tidak	0	Tidak

Lampiran 12. Lanjutan

Penduga Parameter	MLE		Koefisien LASSO	
	Koefisien	Signifikansi	Koefisien	Signifikansi
$\beta_{17(2)}$	-10,765	Tidak	0	Tidak
$\beta_{18(0)}$	-1,188	Tidak	0	Tidak
$\beta_{18(1)}$	-0,862	Tidak	0	Tidak
$\beta_{18(2)}$	11,869	Tidak	0	Tidak
$\beta_{19(0)}$	3,187	Signifikan	0	Tidak
$\beta_{19(1)}$	3,609	Signifikan	0	Tidak
$\beta_{19(2)}$	2,435	Signifikan	0	Tidak
$\beta_{110(0)}$	14,671	Tidak	-0,56	Signifikan
$\beta_{110(1)}$	-0,977	Tidak	1,256	Signifikan
$\beta_{111(0)}$	-16,897	Tidak	0	Tidak
$\beta_{111(1)}$	-16,36	Tidak	0	Tidak
$\beta_{111(2)}$	-13,711	Tidak	0,209	Signifikan
$\beta_{111(3)}$	-15,497	Tidak	-0,473	Signifikan
$\beta_{112(0)}$	2,358	Tidak	0	Tidak
$\beta_{112(1)}$	1,887	Tidak	0	Tidak
$\beta_{112(2)}$	15,505	Tidak	0	Tidak

3. Logit 3(MA)

Penduga Parameter	MLE		Koefisien LASSO	
	Koefisien	Signifikansi	Koefisien	Signifikansi
β_{10}	45,071	Tidak	0,746	Signifikan
β_{11}	0,181	Signifikan	0	Tidak
$\beta_{12(0)}$	-3,173	Signifikan	-2,028	Signifikan
$\beta_{13(0)}$	-16,142	Tidak	0,235	Signifikan
$\beta_{13(1)}$	-15,775	Tidak	0	Tidak
$\beta_{13(2)}$	-3,07	Tidak	0	Tidak
$\beta_{14(0)}$	-13,328	Tidak	0	Tidak
$\beta_{14(1)}$	-13,62	Tidak	0	Tidak
$\beta_{14(2)}$	-13,547	Tidak	0	Tidak

Lampiran 12. Lanjutan

Penduga Parameter	MLE		Koefisien LASSO	
	Koefisien	Signifikansi	Koefisien	Signifikansi
$\beta_{14(3)}$	-12,477	Tidak	0	Tidak
$\beta_{15(0)}$	-10,261	Tidak	0	Tidak
$\beta_{15(1)}$	-8,296	Tidak	0	Tidak
$\beta_{15(2)}$	-10,381	Tidak	0	Tidak
$\beta_{16(0)}$	-8,389	Tidak	0	Tidak
$\beta_{16(1)}$	-8,186	Tidak	0	Tidak
$\beta_{16(2)}$	-8,76	Tidak	0	Tidak
$\beta_{17(0)}$	-11,049	Tidak	0	Tidak
$\beta_{17(1)}$	-10,194	Tidak	0	Tidak
$\beta_{17(2)}$	-10,618	Tidak	0	Tidak
$\beta_{18(0)}$	15,06	Tidak	0	Tidak
$\beta_{18(1)}$	15,178	Tidak	0	Tidak
$\beta_{18(2)}$	28,697	Tidak	0	Tidak
$\beta_{19(0)}$	2,94	Signifikan	0	Tidak
$\beta_{19(1)}$	3,158	Signifikan	0	Tidak
$\beta_{19(2)}$	2,572	Signifikan	0	Tidak
$\beta_{110(0)}$	15,594	Tidak	0	Tidak
$\beta_{110(1)}$	-2,562	Signifikan	0	Tidak
$\beta_{111(0)}$	-16,531	Tidak	0	Tidak
$\beta_{111(1)}$	-16,346	Tidak	0	Tidak
$\beta_{111(2)}$	-14,142	Tidak	0	Tidak
$\beta_{111(3)}$	-14,561	Tidak	0	Tidak
$\beta_{112(0)}$	-1,931	Tidak	0	Tidak
$\beta_{112(1)}$	-2,501	Tidak	0	Tidak
$\beta_{112(2)}$	11,828	Tidak	0	Tidak