

**KETEPATAN KLASIFIKASI PEMILIHAN METODE KONTRASEPSI
DI KOTA SEMARANG MENGGUNAKAN *BOOSTSTRAP AGGREGATING*
REGRESI LOGISTIK MULTINOMIAL**



SKRIPSI

Oleh :

Ahmad Reza Aditya

24010210130055

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

2014

**KETEPATAN KLASIFIKASI PEMILIHAN METODE KONTRASEPSI
DI KOTA SEMARANG MENGGUNAKAN *BOOSTSTRAP*
AGGREGATING REGRESI LOGISTIK MULTINOMIAL**

Disusun Oleh :

Ahmad Reza Aditya

24010210130055

**Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Sains pada Jurusan Statistika**

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

2014

HALAMAN PENGESAHAN I

Judul Skripsi : Ketepatan Klasifikasi Pemilihan Metode Kontrasepsi Di Kota Semarang Menggunakan *Bootstrap Aggregatting* Regresi Logistik Multinomial.

Nama Mahasiswa : Ahmad Reza Aditya

NIM : 24010210130055

Jurusan : Statistika

telah diujikan pada sidang Tugas Akhir tanggal 17 Desember 2014 dan dinyatakan lulus pada tanggal 29 Desember 2014.

Semarang, 29 Desember 2014

Mengetahui

Ketua Jurusan Statistika

FSM Universitas Diponegoro



Despriansanti, M.Si.

NIP. 195709141986032001

Panitia Penguji Tugas Akhir

Ketua

Triastuti Wuryandari, S.Si, M.Si.

NIP. 197109061998032001

HALAMAN PENGESAHAN II

Judul Skripsi : Ketepatan Klasifikasi Pemilihan Metode Kontrasepsi Di Kota Semarang Menggunakan *Bootstrap Aggregating* Regresi Logistik Multinomial.

Nama Mahasiswa : Ahmad Reza Aditya

NIM : 24010210130055

Jurusan : Statistika

telah diujikan pada sidang Tugas Akhir tanggal 17 Desember 2014

Semarang, 29 Desember 2014

Pembimbing I



Dra. Suparti, M.Si.

NIP. 196509131990032001

Pembimbing II



Drs. Sudarno, M.Si.

NIP. 196407091992011001

KATA PENGANTAR

Alhamdulillah, puji syukur penulis panjatkan ke hadirat Allah SWT atas rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir dengan judul **"Ketepatan Klasifikasi Pemilihan Metode Kontrasepsi di Kota Semarang Menggunakan *Bootstrap Aggregating* Regresi Logistik Multinomial"**.

Dalam penulisan laporan tugas akhir ini penulis mengalami banyak hambatan. Namun atas bantuan dari berbagai pihak, laporan tugas akhir ini dapat diselesaikan dengan baik. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih kepada:

1. Ibu Dra. Dwi Ispriyanti, M.Si sebagai Ketua Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro.
2. Ibu Dra. Suparti, M.Si dan Bapak Drs. Sudarno, M.Si selaku dosen pembimbing I dan pembimbing II yang telah memberikan arahan, bimbingan dan motivasi hingga terselesaikannya laporan tugas akhir ini.
3. Bapak/Ibu dosen Jurusan Statistika yang telah memberikan masukan demi perbaikan penulisan laporan tugas akhir ini.
4. Pihak-pihak lain yang tidak dapat penulis sebutkan satu per satu yang telah membantu penulisan laporan tugas akhir ini.

Penulis menyadari bahwa laporan tugas akhir ini masih jauh dari sempurna. Sehingga penulis mengharapkan kritik dan saran demi kesempurnaan penulisan selanjutnya.

Semarang, Desember 2014

Penulis

**KETEPATAN KLASIFIKASI PEMILIHAN METODE KONTRASEPSI
DI KOTA SEMARANG MENGGUNAKAN *BOOSTSTRAP*
AGGREGATING REGRESI LOGISTIK MULTINOMIAL**

ABSTRAK

Pengklasifikasian merupakan salah satu metode statistik dalam pengelompokan suatu data yang disusun secara sistematis. Masalah klasifikasi muncul ketika terdapat sejumlah ukuran yang terdiri dari satu atau beberapa kategori yang tidak dapat diidentifikasi secara langsung tetapi harus menggunakan suatu ukuran. Metode klasifikasi yang biasa digunakan dalam penelitian-penelitian untuk menganalisa suatu masalah atau kejadian adalah metode analisis regresi logistik. Namun metode pengklasifikasian ini memberikan pendugaan parameter yang tidak stabil. Untuk memperoleh parameter yang stabil pada model regresi logistik multinomial digunakan pendekatan bootstrap yaitu metode *bootstrap aggregating (bagging)*. Tujuan dari penelitian ini adalah untuk melihat perbandingan ketepatan klasifikasi model regresi logistik multinomial dan model *bootstrap aggregating* menggunakan data Keluarga Berencana Kota Semarang. Dari hasil *bagging* regresi logistik multinomial diperoleh ketepatan klasifikasi terbesar pada replikasi bootstrap 50 kali yaitu 51%, model ini mampu menurunkan kesalahan klasifikasi hingga 2% dibandingkan dengan model regresi logistik multinomial dengan ketepatan klasifikasi sebesar 49%.

Kata kunci : regresi logistik, bootstrap aggregating, ketepatan klasifikasi

**ACCURACY OF CLASSIFICATION OF CONTRACEPTION METHOD
SELECTION IN SEMARANG USES A BOOTSTRAP AGGREGATING
MULTINOMIAL LOGISTIC REGRESSION**

ABSTRACT

Classification is one of the statistical methods in grouping the data compiled systematically. Classification problem rises when there are a number of measures that consists of one or several categories that can not be identified directly but must use a measure. classification methods commonly used in studies to analyze a problem or event is logistic regression analysis. However, this classification method provides unstable parameter estimation. So to obtain a stable parameter multinomial logistic regression model used bootstrap approach that is bootstrap aggregating (bagging). The purpose of this study was to compare the accuracy of the classification multinomial logistic regression models and bootstrap aggregating model using the data of family planning in Semarang. From the results of bagging multinomial logistic regression obtained classification accuracy in replication bootstrap most 50 times at 51%, this model is able to decrease the classification error of up to 2% compared to the multinomial logistic regression model with a classification accuracy of 49%.

Keywords : logistic regression, bootstrap aggregating, accuracy of classification

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN I	ii
HALAMAN PENGESAHAN II.....	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR LAMPIRAN.....	x
BAB I. PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah.....	4
1.4. Tujuan Penulisan	4
BAB II. TINJAUAN PUSTAKA.....	5
2.1. Regresi Logistik	5
2.2. Regresi Logistik Multinomial	6
2.3. Estimasi Parameter	10
2.4. Pengujian Parameter	13
2.4.1. Pengujian Parameter dengan Uji Likelihood Ratio	13

2.4.2. Pengujian Parameter dengan Uji Wald	14
2.5. Prosedur Klasifikasi	15
2.6. Bagging (<i>Bootstrap Aggregating</i>)	16
BAB III. METODOLOGI PENELITIAN	20
3.1. Sumber Data	20
3.2. Variabel Penelitian	20
3.3. Metode Analisis	22
3.4. Diagram Alir Analisis Data	24
BAB IV. ANALISIS DAN PEMBAHASAN	25
4.1. Regresi Logistik Multinomial	25
4.1.1. Model Awal Regresi Logistik Multinomial	25
4.1.2. Uji Parameter Secara Bersama Model Awal	27
4.1.3. Pengujian Parameter Secara Individu Model Awal	28
4.1.4. Model Kedua Regresi Logistik Multinomial	39
4.1.5. Uji Parameter Secara Bersama Model Kedua	41
4.1.6. Pengujian Parameter Secara Individu Model Kedua	42
4.1.7. Ketepatan Klasifikasi Model Regresi Logistik Multinomial	48
4.2. <i>Bagging</i> Regresi Logistik Multinomial	48
BAB V. KESIMPULAN	52
DAFTAR PUSTAKA	53
LAMPIRAN 54	

DAFTAR TABEL

	Halaman
Tabel 1. Tabel Klasifikasi Untuk 3 Kategori Kelompok	15
Tabel 2. Hasil Uji Wald Variabel X_1 (Model Awal).....	29
Tabel 3. Hasil Uji Wald Variabel X_2 (Model Awal).....	30
Tabel 4. Hasil Uji Wald Variabel X_3 (Model Awal).....	31
Tabel 5. Hasil Uji Wald Variabel X_4 (Model Awal).....	32
Tabel 6. Hasil Uji Wald Variabel X_5 (Model Awal).....	33
Tabel 7. Hasil Uji Wald Variabel X_6 (Model Awal).....	34
Tabel 8. Hasil Uji Wald Variabel X_7 (Model Awal).....	35
Tabel 9. Hasil Uji Wald Variabel X_8 (Model Awal).....	37
Tabel 10. Hasil Uji Wald Variabel X_9 (Model Awal).....	38
Tabel 11. Hasil Uji Wald Variabel X_4 (Model Kedua).....	42
Tabel 12. Hasil Uji Wald Variabel X_7 (Model Kedua).....	44
Tabel 13. Hasil Uji Wald Variabel X_8 (Model Kedua).....	46
Tabel 14. Hasil Uji Wald Variabel X_9 (Model Kedua).....	47
Tabel 15. Ketepatan Klasifikasi Regresi Logistik.....	48
Tabel 16. Hasil Bagging Regresi Logistik Multinomial	49

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data.....	54
Lampiran 2. Regresi Logistik Multinomial.....	61
Lampiran 3. Hasil Ketepatan Klasifikasi Bagging Regresi Logistik Multinomial	74
Lampiran 4. Hasil Estimasi Parameter Model Bagging Regresi Logistik Multinomial	78
Lampiran 5. Sintaks Bagging Regresi Logistik Multinomial pada Matlab	81
Lampiran 6. Tabel Distribusi Chi Square	82

BAB I

PENDAHULUAN

1.1. Latar Belakang

Pengklasifikasian merupakan salah satu metode statistik dalam pengelompokan suatu data yang disusun secara sistematis. Masalah klasifikasi muncul ketika terdapat sejumlah ukuran yang terdiri dari satu atau beberapa kategori yang tidak dapat diidentifikasi secara langsung tetapi harus menggunakan suatu ukuran. Masalah klasifikasi sering dijumpai dalam kehidupan sehari-hari, seperti pengklasifikasian data pada bidang akademik dalam kasus klasifikasi program studi mahasiswa baru, pada bidang sosial dalam kasus klasifikasi tingkat kepuasan kerja karyawan, pada bidang ekonomi dalam kasus klasifikasi tingkat kemiskinan masyarakat, pada bidang perbankan dalam kasus klasifikasi kredit bermasalah, maupun pada bidang lainnya.

Terdapat beberapa metode klasifikasi yang biasa digunakan dalam penelitian-penelitian untuk menganalisa suatu masalah atau kejadian. Salah satunya yaitu metode analisis regresi logistik. Pada regresi logistik akan diperoleh suatu model logistik yang digunakan untuk menjelaskan hubungan antara prediktor dan respon (yang bersifat dikotomis atau ada dua kategori/kelompok), serta untuk mengelompokkan obyek ke dalam salah satu dari dua kategori respon. Dalam perkembangannya, regresi logistik dapat juga digunakan untuk respon kategori lebih dari dua kelompok, yang dikenal dengan regresi logistik polikotomis/multinomial. Regresi logistik merupakan sebuah metode analisis

statistik untuk menggambarkan hubungan antara variabel prediktor dengan variabel respon yang mempunyai dua atau lebih kategori dengan variabel prediktor yang menggunakan skala kategorik maupun interval (Hosmer dan Lemeshow, 1989).

Regresi logistik membentuk persamaan atau fungsi dengan pendekatan *Maximum Likelihood*, yang memaksimalkan peluang pengklasifikasian objek yang diamati menjadi kategori yang sesuai kemudian mengubahnya menjadi koefisien regresi yang sederhana. Dalam hal ini regresi logistik tidak memodelkan secara langsung variabel respon dengan variabel prediktor-prediktornya, melainkan melalui transformasi variabel respon ke variabel logit yang merupakan *natural log* dari *Odds Rasio*. Metode analisis regresi logistik dalam beberapa literatur klasifikasi sering disebut sebagai model klasik/parametrik. Metode pengklasifikasian ini memberikan pendugaan parameter yang tidak stabil, artinya jika terdapat perubahan dalam data set menyebabkan perubahan yang signifikan pada model (Breiman, 1994). Untuk memperoleh parameter yang stabil pada model regresi logistik multinomial digunakan pendekatan bootstrap yaitu metode *Bootstrap Aggregating (Bagging)*. Bagging diperkenalkan oleh Breiman (1994) adalah metode untuk memperbaiki kekuatan prediksi dari beberapa penduga atau algoritma tertentu seperti Regresi atau pohon klasifikasi. *Bagging* regresi logistik bekerja dengan cara mengambil sampel berukuran n dari data asli berukuran n secara acak dengan pengembalian sebanyak B kali dan memodelkan Regresi logistik untuk setiap sampel pada setiap B pengulangan.

Metode yang dinyatakan terbaik dalam perbandingan antara regresi logistik multinomial dengan *Bagging* Regresi logistik multinomial biasanya

adalah metode yang memiliki tingkat kesalahan klasifikasi (*missclassified*) lebih kecil. Tingkat kesalahan klasifikasi dapat diketahui dari hasil ketepatan prediksi masing-masing metode yang dibandingkan dengan data aktualnya. Untuk lebih jelasnya bagaimana Regresi Logistik Multinomial dan *Bootstrap Aggregating* Regresi Logistik Multinomial bekerja dan metode mana yang lebih baik dalam memprediksi maka penulis mencoba mengaplikasikannya pada data keluarga di kota Semarang 2013. Data tersebut merupakan data sekunder yang diambil dari pemutakhiran data keluarga Januari 2013-Januari 2014 (MDK) yang dilakukan oleh BKKBN Jawa Tengah.

Variabel yang digunakan adalah faktor yang mempengaruhi pemilihan alat kontrasepsi pada wanita, antara lain: umur istri, umur suami, umur anak terakhir, jumlah anak hidup, tingkat pendidikan formal suami dan istri, jenis pekerjaan suami, jenis pekerjaan istri, dan tingkat kesejahteraan keluarga. Faktor-faktor tersebut akan mempengaruhi penggunaan alat kontrasepsi, variabel diukur dalam skala kategori, sedangkan alat kontrasepsi wanita, yaitu IUD, Kondom, Suntikan, MOW, MOP, Implant, dan Pil sebagai variabel yang dipengaruhi juga bersifat kategori (Sulistio, 2010).

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka beberapa masalah yang akan dibahas dalam penelitian ini adalah sebagai berikut

1. Bagaimana langkah-langkah pemodelan regresi logistik multinomial?
2. Bagaimana bentuk model regresi logistik multinomial dalam pemilihan metode kontrasepsi di kota Semarang?

3. Bagaimana proses *Bootstrap Aggregating* pada model regresi logistik multinomial dalam pemilihan metode kontrasepsi di kota Semarang?
4. Bagaimana perbandingan ketepatan klasifikasi model Regresi Logistik Multinomial dan model *Bootstrap Aggregating* Regresi Logistik Multinomial dalam pemilihan metode kontrasepsi di kota Semarang?

1.3. Batasan Masalah

Ruang lingkup penelitian ini adalah mengkaji dan membandingkan model Regresi Logistik Multinomial dan *Bootstrap Aggregating* Regresi Logistik Multinomial dalam menganalisis masalah klasifikasi pemilihan metode kontrasepsi di kota Semarang dari pemutakhiran data keluarga Januari 2013-Januari 2014 (MDK) yang dilakukan oleh BKKBN Jawa Tengah.

1.4. Tujuan Penulisan

Tujuan dari penulisan ini adalah

1. Mengetahui bentuk model dan ketepatan klasifikasi Regresi Logistik Multinomial dalam pemilihan metode kontrasepsi di kota Semarang.
2. Mengetahui proses *Bootstrap Aggregating* pada model Regresi Logistik Multinomial dan ketepatan klasifikasi dalam pemilihan metode kontrasepsi di kota Semarang.
3. Membandingkan ketepatan klasifikasi model Regresi Logistik Multinomial dan model *Bootstrap Aggregating* Regresi Logistik Multinomial dalam pemilihan metode kontrasepsi di kota Semarang.