

**PERBANDINGAN ANALISIS DISKRIMINAN LINIER  
KLASIK DAN ANALISIS DISKRIMINAN LINIER ROBUST  
UNTUK PENGLASIFIKASIAN KESEJAHTERAAN  
MASYARAKAT KABUPATEN/KOTA DI JAWA TENGAH**



**SKRIPSI**

**Disusun Oleh:  
Ana Kartikawati  
NIM. J2E009024**

**JURUSAN STATISTIKA  
FAKULTAS SAINS DAN MATEMATIKA  
UNIVERSITAS DIPONEGORO  
SEMARANG  
2013**

**PERBANDINGAN ANALISIS DISKRIMINAN LINIER KLASIK DAN  
ANALISIS DISKRIMINAN LINIER ROBUST UNTUK  
PENGKLASIFIKASIAN KESEJAHTERAAN MASYARAKAT  
KABUPATEN/KOTA DI JAWA TENGAH**

**Disusun Oleh :**

**ANA KARTIKAWATI**

**J2E009024**

Skripsi

Sebagai Salah Satu Syarat untuk Memperoleh Gelar

Sarjana Sains pada Jurusan Statistika

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

2013

## HALAMAN PENGESAHAN I

Judul : **Perbandingan Analisis Diskriminan Linier Klasik dan Analisis Diskriminan Linier Robust untuk Pengklasifikasian Kesejahteraan Masyarakat Kabupaten/Kota di Jawa Tengah**

Nama : Ana Kartikawati

NIM : J2E009024

Telah diujikan pada sidang Tugas Akhir tanggal 24 Juni 2013 dan dinyatakan lulus pada tanggal 5 Juli 2013.

Semarang, 5 Juli 2013

Mengetahui,

Ketua Jurusan Statistika

Fakultas Sains dan Matematika UNDIP,



Dra. Dwi Ispriyanti, M.Si.  
NIP. 195709141986032001

Panitia Penguji Ujian Tugas Akhir

Ketua,



Dra. Suparti, M.Si.  
NIP. 196509131990032001

**HALAMAN PENGESAHAN II**

Judul : **Perbandingan Analisis Diskriminan Linier Klasik dan Analisis Diskriminan Linier Robust untuk Pengklasifikasian Kesejahteraan Masyarakat Kabupaten/Kota di Jawa Tengah**

Nama : Ana Kartikawati

NIM : J2E009024

Telah diujikan pada sidang Tugas Akhir tanggal 24 Juni 2013.

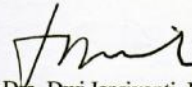
Semarang, 5 Juli 2013

Pembimbing I



Moch. Abdul Mukid, S. Si., M. Si.  
NIP. 197808172005011001

Pembimbing II



Dra. Dwi Ispriyanti, M.Si.  
NIP. 195709141986032001

## KATA PENGANTAR

Puji syukur penulis ucapkan atas kehadiran Allah SWT yang telah melimpahkan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi berjudul “Perbandingan Analisis Diskriminan Linier Klasik dan Analisis Diskriminan Linier Robust untuk Pengklasifikasian Kesejahteraan Masyarakat Kabupaten/Kota di Jawa Tengah”. Skripsi ini menerangkan suatu perbandingan metode analisis data yang dapat mengklasifikasi tingkat kesejahteraan masyarakat kabupaten/kota di Jawa Tengah dengan menggunakan penaksir klasik dan penaksir *robust*.

Keberhasilan penulis dalam menyelesaikan skripsi ini tentu bukan hanya usaha dan doa penulis seorang, tetapi banyak pihak yang telah berjasa dan membantu dalam penyusunan skripsi ini. Oleh karena itu, pada kesempatan ini penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Dra. Hj. Dwi Ispriyanti, M.Si. selaku Ketua Jurusan Statistika FSM Universitas Diponegoro Semarang
2. Moch. Abdul Mukid, S.Si., M.Si. selaku dosen pembimbing I dan Dra. Hj. Dwi Ispriyanti, M.Si. selaku dosen pembimbing II yang dengan penuh kesabaran telah memberikan bimbingan, pengarahan dan petunjuk dalam penulisan skripsi ini
3. Seluruh Bapak dan Ibu Dosen Jurusan Statistika, FSM Universitas Diponegoro yang telah memberikan ilmu yang sangat bermanfaat
4. Kedua orang tua penulis, H. Riyatno, ST dan Hj. Karniti yang senantiasa memberikan kasih sayang yang tulus, doa dan semangat tiada henti

5. Semua pihak yang telah memberikan bantuan, semangat dan doa sehingga penulis dapat menyelesaikan penulisan skripsi ini

Penulis menyadari bahwa penulisan skripsi ini belumlah sempurna. Oleh karena itu, kritik dan saran yang bersifat membangun sangat penulis harapkan. Penulis berharap semoga skripsi ini dapat bermanfaat bagi penulis pada khususnya maupun pembaca pada umumnya.

Semarang, Juli 2013

Penulis

## ABSTRAK

Analisis diskriminan merupakan suatu metode statistika yang digunakan untuk mengklasifikasikan suatu individu atau objek ke dalam suatu kelompok yang telah ditentukan sebelumnya berdasarkan variabel-variabel bebasnya. Analisis diskriminan yang biasa digunakan ialah analisis diskriminan klasik yang meliputi analisis diskriminan linier klasik dan analisis diskriminan kuadrat klasik. Dalam analisis diskriminan linier klasik terdapat dua asumsi yang harus dipenuhi, yaitu variabel bebas berdistribusi normal multivariat dan matriks kovariansi dari kedua kelompok yang diamati adalah sama. Analisis diskriminan klasik tidak dapat bekerja dengan baik jika data yang dianalisis mengandung banyak pencilan. Agar analisis diskriminan tetap optimal dalam pengklasifikasian meskipun dalam kondisi data yang mengandung banyak pencilan maka diperlukan suatu penaksir yang *robust*. Melalui analisis diskriminan *robust* diharapkan diperoleh ketepatan klasifikasi yang tinggi untuk data yang mengandung banyak pencilan. Penaksir *fast-MCD* merupakan salah satu penaksir *robust* yang bertujuan untuk mendapatkan determinan matriks kovariansi terkecil. Analisis diskriminan linier *robust* dengan metode *fast-MCD* pada penulisan tugas akhir ini diterapkan untuk menentukan status kesejahteraan masyarakat kabupaten/kota di Jawa Tengah. Total proporsi ketepatan pengklasifikasian dengan metode analisis diskriminan linier *robust* pada data kesejahteraan masyarakat kabupaten/kota di Jawa Tengah yaitu sebanyak 77,14 persen sama dengan metode analisis diskriminan linier klasik yang juga sebesar 77,14 persen. Hal ini dikarenakan jumlah pencilan yang terlalu kecil pada data kesejahteraan masyarakat kabupaten/kota di Jawa Tengah.

Kata Kunci : Kesejahteraan, Pencilan, Analisis Diskriminan Linier *Robust*

## ABSTRACT

Discriminant analysis is a statistics method which is used to classify an individual or object into certain group which has determined based on its independent variables. Discriminant analysis that commonly used is classical discriminant analysis which consist of classical linear discriminant analysis and classical quadratic discriminant analysis. In classical linear discriminant analysis there are two assumptions to be fulfilled i.e. independent variables have to be normal multivariate distributed and the covariance matrix from the two observed objects should be the same. Classical discriminant analysis cannot work properly if the data which being analyzed consists of many outliers. In order to make discriminant analysis works optimally within the classification though in the condition of data which contains of many outliers, *robust* estimator is needed. The *robust* discriminant analysis is used to get the high classification accuracy for data which contains of many outliers. *Fast-MCD* estimator is one of the *robust* estimators which is aimed to get the smallest determinant of covariance matrices. The *robust* linear discriminant analysis with *fast-MCD* method in this graduating paper is implemented to determine the prosperity status of the people in the regencies or towns in Central Java. The total proportion of classification accuracy using *robust* linear discriminant analysis method on the data of Central Java people prosperity is 77.14 percent. It is equal with the result from classic linear discriminant analysis which is also 77.14 percent. It is caused by the few amount of outlier on the data of Central Java people prosperity.

Keywords: Prosperity, Outlier, *Robust* Linear Discriminant Analysis



## DAFTAR ISI

	<b>Halaman</b>
<b>HALAMAN JUDUL</b> .....	i
<b>HALAMAN PENGESAHAN</b> .....	ii
<b>KATA PENGANTAR</b> .....	iv
<b>ABSTRAK</b> .....	vi
<b>ABSTRACT</b> .....	vii
<b>DAFTAR ISI</b> .....	viii
<b>DAFTAR GAMBAR</b> .....	xii
<b>DAFTAR TABEL</b> .....	xiii
<b>DAFTAR LAMPIRAN</b> .....	xiv
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang.....	1
1.2 Tujuan Penulisan.....	3
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Kesejahteraan Masyarakat.....	4
2.1.1 Upah Minimum Kabupaten/Kota.....	6
2.1.2 Tingkat Pengangguran.....	7
2.1.3 Tingkat Pertumbuhan Ekonomi.....	7
2.1.4 Tingkat Inflasi.....	7
2.1.5. Jumlah Puskesmas.....	8
2.2 Distribusi Normal Multivariat .....	8

2.3	Kesamaan Matriks Varian Kovarian.....	9
2.4	Pendeteksian Pencilan.....	10
2.5	Analisis Diskriminan .....	11
2.5.1	Pemilihan Variabel Bebas.....	12
2.5.2	Pendekatan Fisher untuk Klasifikasi dengan Dua Populasi.....	13
2.5.3	Perhitungan Kesalahan Pengklasifikasian Analisis Diskriminan .....	17
2.5.4	Menilai Keakuratan Prediksi Keanggotaan Kelompok.....	19
2.6	Metode Penaksir <i>MCD</i> .....	19
2.6.1	Metode <i>fast-MCD</i> .....	20
2.7	Analisis Diskriminan Linier Robust .....	21

### **BAB III METODOLOGI PENELITIAN**

3.1	Sumber Data .....	23
3.2	Variabel Penelitian .....	23
3.3	Langkah-langkah Analisis .....	23
3.4	Diagram Alir Analisis Data .....	25

### **BAB IV HASIL DAN PEMBAHASAN**

4.1	Data Tingkat Kesejahteraan Masyarakat Kabupaten/Kota di Jawa Tengah.....	28
4.1.1	Upah Minimum Kabupaten/Kota.....	28
4.1.2	Tingkat Pengangguran.....	29
4.1.3	Tingkat Pertumbuhan Ekonomi.....	30
4.1.4	Tingkat Inflasi.....	31
4.1.5	Jumlah Puskesmas.....	32

4.2	Pengujian Asumsi Analisis Diskriminan.....	33
4.2.1	Uji Normal Multivariat.....	33
4.2.2	Uji Kesamaan Matriks Varian Kovarian.....	35
4.3	Pendeteksian Pencilan.....	36
4.4	Pemilihan Variabel Bebas.....	37
4.4.1	Variabel Upah Minimum Kabupaten/Kota.....	37
4.4.2	Variabel Tingkat Pengangguran.....	37
4.4.3	Variabel Tingkat Pertumbuhan Ekonomi.....	38
4.4.4	Variabel Tingkat Inflasi.....	39
4.4.5	Variabel Jumlah Puskesmas.....	39
4.5	Analisis Diskriminan Linier Klasik.....	40
4.5.1	Penaksiran Parameter.....	40
4.5.2	Fungsi Diskriminan.....	42
4.5.3	Pengklasifikasian Pengamatan.....	43
4.5.4	Perhitungan Kesalahan Pengklasifikasian.....	44
4.5.5	Keakuratan Prediksi Keanggotaan Kelompok.....	46
4.6	Analisis Diskriminan Linier <i>Robust</i> .....	46
4.6.1	Metode <i>fast-MCD</i> .....	46
4.6.2	Fungsi Diskriminan.....	48
4.6.3	Pengklasifikasian Pengamatan.....	49
4.6.4	Perhitungan Kesalahan Pengklasifikasian.....	50
4.6.5	Keakuratan Prediksi Keanggotaan Kelompok.....	52
4.7	Perbandingan Hasil Analisis Diskriminan Linier Klasik dengan Analisis Diskriminan Linier <i>Robust</i> .....	52

**BAB V KESIMPULAN**

5.1 Kesimpulan .....	54
5.2 Saran .....	54
<b>DAFTAR PUSTAKA .....</b>	<b>56</b>
<b>LAMPIRAN.....</b>	<b>58</b>

## DAFTAR GAMBAR

	<b>Halaman</b>
<b>Gambar 3.1.</b> Diagram Alir Analisis Data Penelitian .....	26
<b>Gambar 4.1.</b> Perbandingan Upah Minimum Kabupaten/Kota dengan Status Tingkat Kesejahteraan Tinggi dan Kesejahteraan Rendah.....	29
<b>Gambar 4.2.</b> Perbandingan Tingkat Pengangguran Kabupaten/Kota dengan Status Kesejahteraan Tinggi dan Kesejahteraan Rendah.....	30
<b>Gambar 4.3.</b> Perbandingan Tingkat Pertumbuhan Ekonomi Kabupaten/ Kota dengan Status Kesejahteraan Tinggi dan Kesejahteraan Rendah .....	31
<b>Gambar 4.4.</b> Perbandingan Tingkat Inflasi Kabupaten/Kota dengan Status Kesejahteraan Tinggi dan Kesejahteraan Rendah.....	32
<b>Gambar 4.5.</b> Perbandingan Jumlah Puskesmas Kabupaten/Kota dengan Status Kesejahteraan Tinggi dan Kesejahteraan Rendah.....	33

## DAFTAR TABEL

	<b>Halaman</b>
<b>Tabel 2.1</b> Klasifikasi Dua Kelompok.....	18
<b>Tabel 4.1</b> Data Status Kesejahteraan Kabupaten/Kota di Jawa Tengah.....	34
<b>Tabel 4.2</b> Skor Diskriminan dan Hasil Pengklasifikasian Analisis Diskriminan Linier Klasik.....	45
<b>Tabel 4.3</b> Kesalahan Pengklasifikasian Analisis Diskriminan Linier Klasik..	44
<b>Tabel 4.4</b> Skor Diskriminan dan Hasil Pengklasifikasian Analisis Diskriminan Linier <i>Robust</i> .....	51
<b>Tabel 4.5</b> Kesalahan Pengklasifikasian Analisis Diskriminan Linier <i>Robust</i>	50
<b>Tabel 4.6</b> Perbandingan Kesalahan Pengklasifikasian.....	53

## DAFTAR LAMPIRAN

	<b>Halaman</b>
Lampiran I	Data Tingkat Kesejahteraan Masyarakat Kabupaten/Kota di Jawa Tengah..... 58
Lampiran II	Data Tingkat Kesejahteraan Masyarakat Kabupaten/Kota di Jawa Tengah Setelah Distandarisasi..... 59
Lampiran III	Pengujian Normal Multivariat Menggunakan R 2.14.0..... 60
Lampiran IV	Pengujian Kesamaan Matriks Varian Kovarian Menggunakan SPSS 13.00 For Windows..... 61
Lampiran V	Perhitungan untuk Mendeteksi Pencilan Menggunakan Minitab 14..... 62
Lampiran VI	Hasil Pendeteksian Pencilan..... 64
Lampiran VII	Perhitungan untuk Mendapatkan Penaksir Klasik dengan Menggunakan Minitab 14..... 65
Lampiran VIII	Paket Program untuk Mendapatkan Penaksir <i>fast-MCD</i> dengan Menggunakan R 2.14.0. .... 67
Lampiran IX	Tabel Distribusi Chi-Kuadrat..... 68
Lampiran X	Tabel Distribusi t..... 69

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Keadaan masyarakat Indonesia saat ini sungguh memprihatinkan. Hal ini dapat terlihat dari masih banyaknya masyarakat miskin di sejumlah daerah. Padahal kesejahteraan ekonomi sangat dibutuhkan oleh setiap masyarakat. Oleh karena itu, sudah seharusnya dilaksanakan pemerataan pendapatan agar kesejahteraan dapat dirasakan oleh setiap masyarakat. Berdasarkan hal tersebut, pemerintah sebaiknya perlu menentukan status tingkat kesejahteraan masyarakat suatu daerah agar dapat dilaksanakan pembangunan pada daerah-daerah yang tingkat kesejahteraannya masih terbilang rendah. Salah satu alat untuk menentukan status tingkat kesejahteraan masyarakat di suatu wilayah ialah pola pengeluaran konsumsi masyarakat (BPS, 2011). Dalam ilmu statistik terdapat berbagai macam metode statistika untuk menganalisis berbagai macam permasalahan seperti menentukan status kesejahteraan masyarakat tinggi atau rendah. Salah satu metode yang dapat digunakan untuk permasalahan tersebut ialah analisis diskriminan.

Analisis diskriminan merupakan suatu metode statistika yang digunakan untuk mengklasifikasikan suatu individu atau objek ke dalam suatu kelompok yang telah ditentukan sebelumnya berdasarkan variabel-variabel bebasnya (Dillon dan Goldstein, 1984). Analisis diskriminan yang biasa digunakan ialah analisis diskriminan klasik yang meliputi analisis diskriminan linier klasik dan analisis diskriminan kuadrat klasik. Dalam analisis diskriminan linier klasik



terdapat dua asumsi yang harus dipenuhi, yaitu variabel bebas berdistribusi normal multivariat dan matriks kovariansi dari kedua kelompok data yang diamati adalah sama, sedangkan asumsi analisis diskriminan kuadratik klasik yaitu variabel bebas berdistribusi normal multivariat dan matriks kovariansi dari kedua kelompok data yang diamati adalah berbeda.

Analisis diskriminan klasik tidak dapat bekerja dengan baik jika data yang dianalisis mengandung pencilan (An dan Jin, 2011). Hal ini dikarenakan rata-rata sampel dan matriks varian kovarian sampel sangat sensitif terhadap pencilan dan berpengaruh pada keakuratan hasil pengklasifikasian analisis diskriminan. Agar analisis diskriminan tetap optimal dalam pengklasifikasian meskipun dalam kondisi data yang mengandung pencilan maka diperlukan suatu penaksir yang *robust*. Analisis diskriminan yang menggunakan penaksir *robust* selanjutnya disebut sebagai analisis diskriminan *robust*.

Beberapa metode pada analisis diskriminan *robust* diantaranya metode *the minimum volume ellipsoid (MVE)*, *fast minimum covariance determinant (fast-MCD)* dan *M-estimator* (Rousseuw dan Driessen, 1999). Metode *fast-MCD* belum begitu banyak dipakai dalam mengatasi data pencilan. Oleh karena itu, penaksir *robust* yang digunakan pada penulisan tugas akhir ini ialah penaksir *fast-MCD*. Selain itu, analisis diskriminan yang digunakan pada penulisan tugas akhir ini ialah analisis diskriminan linier. Berdasarkan hal tersebut, dalam penulisan tugas akhir ini analisis diskriminan linier *robust* dengan metode *fast-MCD* diterapkan untuk menentukan status kesejahteraan masyarakat kabupaten/kota di Jawa Tengah.

Dalam penulisan tugas akhir ini, permasalahan yang dibahas yaitu pengklasifikasian data tingkat kesejahteraan masyarakat kabupaten/kota di Jawa Tengah berdasarkan fungsi diskriminan linier klasik dua kelompok dan fungsi diskriminan linier *robust* dua kelompok serta perbandingan hasil analisis diskriminan linier klasik dan hasil analisis diskriminan linier *robust* pada data tingkat kesejahteraan masyarakat kabupaten/kota di Jawa Tengah. Masalah dibatasi hanya pada dua kelompok populasi dan lima variabel bebas yang digunakan dalam menentukan status kesejahteraan masyarakat kabupaten/kota di Jawa Tengah, yaitu upah minimum kabupaten/kota (UMK), tingkat pengangguran, tingkat pertumbuhan ekonomi, tingkat inflasi dan jumlah puskesmas. Tidak hanya itu, masalah juga dibatasi pada metode yang digunakan dalam analisis diskriminan linier *robust* ini yaitu dengan menggunakan metode *fast-MCD*. Pada akhirnya, penelitian ini diharapkan dapat bermanfaat untuk mengetahui variabel apa saja yang dapat dijadikan dasar untuk mengklasifikasi tingkat kesejahteraan masyarakat kabupaten/kota di Jawa Tengah.

## **1.2 Tujuan Penulisan**

Tujuan dari penelitian dalam penulisan tugas akhir ini adalah

1. Mengklasifikasikan data tingkat kesejahteraan masyarakat kabupaten/kota di Jawa Tengah berdasarkan fungsi diskriminan linier klasik dua kelompok
2. Mengklasifikasikan data tingkat kesejahteraan masyarakat kabupaten/kota di Jawa Tengah berdasarkan fungsi diskriminan linier *robust* dua kelompok
3. Membandingkan hasil analisis diskriminan linier klasik dengan hasil analisis diskriminan linier *robust* pada data tingkat kesejahteraan masyarakat kabupaten/kota di Jawa Tengah