



# SEMINAR NASIONAL SAINS MIPA DAN APLIKASINYA

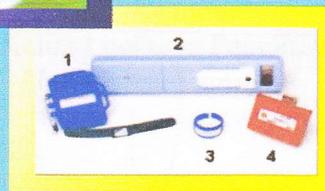
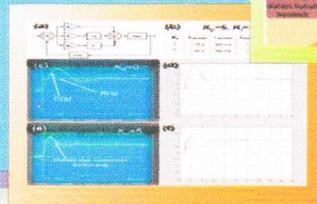
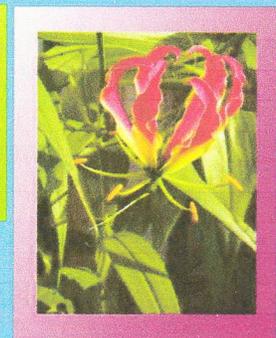
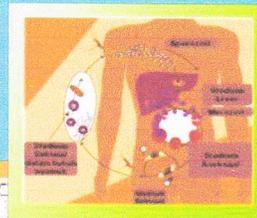
Bandar Lampung, 8 - 9 Desember 2010



Tema :  
"Membangun Jejaring Pengembang  
Dan Pengguna SAINS dan Matematika"

## PROSIDING

ISSN: 2086-2342



Vol.2 Tahun 2010  
Buku 2

---

**PROSIDING  
SEMINAR NASIONAL SAINS MIPA DAN APLIKASINYA 2010  
(SN SMAP 10)**

**TEAM PENYERTING**  
Nurcholis Agumata, Ph.D.  
Wahyuni, Ph.D.  
Sriani Samudra, Ph.D.  
Mulyono, Ph.D.  
Wardana, Ph.D.  
Dan Susanto Pratiwi, M.S.  
Wahid Lase Tama Salsary, M.S.

**PEMBAKIT**  
PILAKA Mardiana dan Abu Pengabdian Hari, Universitas Lampung

**LOKASI KEDIRYAN**  
Gedung Departemen Matematika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Jl. S. Soedjatmoko No. 1  
Telp./Faks: 440.741.2000  
Email: [mat@unila.ac.id](mailto:mat@unila.ac.id)  
Email: [snsmap@unila.ac.id](mailto:snsmap@unila.ac.id)



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
DESEMBER 2010**

# **Prosiding** Seminar Nasional Sains MIPA dan Aplikasinya Tahun 2010 FMIPA Universitas Lampung

## **TEAM PENYUNTING :**

Rochmah Agustrina, Ph.D.  
Wasinton, Ph.D.  
Simon Sembiring, Ph.D.  
Mulyono, Ph.D.  
Warsono, Ph.D.  
Dian Septiani Pratama, M.Si.  
Ni Luh Gede Ratna Juliasih, M.Si.

## **PENERBIT**

Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung

## **ALAMAT REDAKSI**

Gedung Dekanat Lantai 4  
Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung  
Jl. S. Brodjonegoro No. 1, Bandar Lampung 35145  
Telp./Fax: +62-721-704625;  
<http://fmipa.unila.ac.id/>  
E-mail: [seminar-smap@unila.ac.id](mailto:seminar-smap@unila.ac.id)

Prosiding Seminar Nasional  
Sains MIPA dan Aplikasinya FMIPA UNILA:  
penyunting, Rochmah Agustrina[et al.]  
Desember 2010 / — Bandar Lampung  
ix + 248 hlm.; 21 x 29,7 cm

**ISSN 2086-2342**  
(Terbit satu kali setahun)



## KATA PENGANTAR

Assalamu'alaikum wa Rohmatullahi wa Barokatuhu.

Alhamdulillah dengan perkenan-Nya lah, maka Prosiding Seminar Nasional Sains MIPA dan Aplikasinya tahun 2010 (SN SMAP 10) 8 – 9 Desember 2010 dengan tema: " Membangun Jejaring Pengembang dan Pengguna Sains dan Matematika", telah dapat kami selesaikan. Kegiatan seminar ini merupakan salah satu rangkaian kegiatan dalam rangka Dies Natalis FMIPA UNILA, yang diagendakan dilakukan secara rutin tahunan.

Segenap panitia mengucapkan terima kasih kepada Rektor UNILA Bapak Prof. Dr. Ir. Sugeng P Harianto, M.S. dan Dekan Fakultas MIPA Bapak Dr. Sutyarso, M.Biomed. yang telah memfasilitasi berlangsungnya kegiatan ini. Demikian pula kepada para Keynote Speakers: Ir. Edi Yanto, M.Si., Kepala Bappeda Provinsi Lampung; Prof. Dr. Ir. Suprpto DEA., Ketua Tim Ahli Pengembangan HKI, DP2M- Ditjen Dikti, Kemendiknas dan Dosen Jurusan Teknik Kimia FTI - ITS; dan Bapak Dr. Sutyarso, M.Biomed., Dekan FMIPA Universitas Lampung, yang telah berkenan memberikan materi pada kegiatan ini.

Kami juga menyampaikan penghargaan dan terima kasih atas apresiasi rekan-rekan akademisi dan peneliti untuk berkenan mempresentasikan hasil penelitiannya dalam kegiatan Seminar Nasional ini. Seminar ini diikuti oleh berbagai kelompok Sains MIPA dan aplikasinya dalam kategori kelompok ilmu Matematika, Fisika, Biologi, dan Kimia. Jumlah makalah yang dipresentasikan dalam kegiatan ini sebanyak 84 makalah dan yang masuk dalam prosiding ini adalah sebanyak 83 makalah. Akhir kata, kami sampaikan terima kasih kepada semua pihak yang telah mendukung penyusunan prosiding kegiatan seminar ini. Dalam kesempatan ini, kami juga memohon maaf apabila ada hal hal yang kurang berkenan selama pelaksanaan kegiatan seminar ataupun dalam penyusunan prosiding seminar ini. Akhir kata mari kita bersama mendukung upaya peningkatan daya saing bangsa melalui karya nyata dalam bidang Sains MIPA dan Aplikasinya.

Wassalamu'alaikum wa Rohmatullahi wa Barokatuhu.

Tim Penyunting .

## DAFTAR ISI

<b>Kelompok Fisika</b>	halaman
KALIBRASI EFISIENSI SISTEM PENCACAH KAMAR PENGIONAN 4 $\pi$ - $\gamma$ MERLIN GERIN SEBAGAI ALAT STANDAR SEKUNDER PENGUKURAN AKTIVITAS MENGGUNAKAN SUMBER STANDAR CAIR Hermawan Candra dan Holnisar	1-9
KALIBRASI DETEKTOR KEPING SEJAJAR ROOS TERHADAP DETEKTOR SILINDRIS UNTUK BERKAS SINAR GAMMA Co-60 DAN ENERGI NOMINAL BERKAS ELEKTRON 20 MeV C. Tuti Budiantari dan Nurman R	10-18
ANALISIS PERFORMANSI PENGONTROL ON-OFF PADA RESPON SISTEM KENDALI UMPAN BALIK Darmawan Hidayat, Meutia Rahmatika, Tuti Aryati Demen, Jajat Yuda Mindara	19-24
KALIBRASI <i>IN SITU</i> MONITOR AREA DI PUSAT <i>PET-CT SCAN</i> DAN <i>CYCLOTRON</i> RS GADING PLUIT, JAKARTA Nurman Rajagukguk dan C Tuti Budiantari	25-28
APLIKASI TURUNAN PERTAMA DAN KEDUA VERTIKAL PADA ANALISA DATA GRAVITASI DAN GEOMAGNET Syamsu Rosid dan Erlangga Harhale	29-35
KARAKTERISTIK KERAMIK <i>MULLITE</i> DARI SILIKA SEKAM PADI AKIBAT PERLAKUAN KALSINASI Sherly Nevivilanti, Frissilla Venia Wiranti dan Simon Sembiring	36-40
PENENTUAN DAERAH KERJA OPTIMAL DARI IDENTIFIKASI DETEKTOR HPRID 75023 MENGGUNAKAN Ra-226, Co-60, Cs-137 DAN I-131 Wijono, Gatot Wurdianto dan Holnisar	41-46
ANALISIS RESPON SISTEM KENDALI DENGAN PENGONTROL PID PADA MODEL FISIS ELEKTRONIK Tuti Aryati Demen, Eppstian Syah As'ari, Darmawan Hidayat, Bambang Mukti Wibawa	47-53
KARAKTERISTIK STRUKTUR KERAMIK BOROSILIKAT BERBASIS SILIKA SEKAM PADI AKIBAT PERLAKUAN SINTERING Simon Sembiring	54-59
SIMULASI DAN PREDIKSI VARIABILITAS IKLIM DAERAH LAMPUNG BERBASIS MODEL REGIONAL DARLAM Bambang Siswanto	60-69
MODEL PEMBELAJARAN FISIKA BERBASIS <i>PROBLEM SOLVING</i> UNTUK MENINGKATKAN KEMAMPUAN METAKOGNISI CALON GURU Mariati Purnama Simanjuntak, Agus Setiawan, Andi Suhandi	70-77

PEMODELAN INVERSI 3D ANOMALI GAYABERAT UNTUK MENGETAHUI KEBERADAAN KANTONG MAGMA GUNUNG MERAPI Muh Sarkowi	78-84
STUDI DAN ANALISA NILAI PERMEABILITAS DARI UJI PERMEABILITAS SKALA LAPANGAN DAN SKALA LABORATORIUM Ir. Setyanto, M.T dan Dr.Ir.Lusmelia Afriani	85-92
VARIASI KADAR KCI DALAM PROSES PELELEHAN PADA PEMBENTUKAN FASE BAHAN SUPERKONDUKTOR BSCCO-2223 Suprihatin	93-99
UJI TAK RUSAK BAHAN KONDUKTOR DENGAN METODE PEMINDAI PANAS MENGGUNAKAN MIKROKONTROLER ATMEGA 16 Warsito, Muhammad Wahyu Rizal, Sri Wahyu Suciwati	100-107
APLIKASI TRANSFORMASI HARTLEY PADA KONTINUASI DATA GRAVITASI DAN MAGNETIK Syamsu Rosid dan Benny Irawan	108-115
PENDUGAAN PENCEMARAN AIR TANAH SEKITAR ALIRAN SUNGAI DI KAWASAN INDUSTRI RANCAEKEK BERDASARKAN DISTRIBUSI NILAI RESISTIVITAS HASIL PENGUKURAN RESISTIVITAS-DC Asep Harja	116-127
DELINIASI LAPISAN TANAH PENYEBAB AMBLESAN BADAN JALAN KM- 79/80 DESA HATTA, KEC. BAKAUHENI DENGAN GEOLISTRIK DIPOLE-DIPOLE Ahmad Zaenudin	128-133

**Kelompok Matematika**

SYARAT PERLU DAN CUKUP SUATU MODUL MEMPUNYAI AMPLOP PROYEKTIF DI DALAM $\sigma[M]$ Fitriani	134-138
APLIKASI MODEL TRANSPORTASI DALAM MENGOPTIMASI BIAYA PELAYANAN JASA KEGIATAN PENERBANGAN DOMESTIK PESAWAT ANGKUTAN PENUMPANG Sugandi Yahdin , Ning Eliyati , Subrina	139-146
STUDI PEMODELAN STOKASTIK CURAH HUJAN HARIAN DARI DATA CURAH HUJAN STASIUN PURAJAYA Ahmad Zakaria	147-155
MODEL ANNUITAS SYARIAH DENGAN PENDEKATAN FUZZY Endang Sri Kresnawati <sup>1</sup>	156-161
PEMODELAN PRODUK ASURANSI CACAT DENGAN SUKU BUNGA SEMI MARKOV NONHOMOGEN DISKRIT Alfensi Faruk	162-167

METODE K-RATAAN TERPANGKAS ( <i>TRIMMED K-MEANS</i> ) DALAM MENGATASI PENCILAN PADA DATA BERGEROMBOL Bunga Ayu Ningrum Hamka	168-176
METODE PENENTUAN MODEL REGRESI BERGANDA TERBAIK MENGUNAKAN MODEL FRAKSI Herlina Hanum	177-183
ANALISIS DISKRIMINAN LINIER ROBUST MENGGUNAKAN METODE VOLUME ELLIPSOID MINIMUM Khoirin Nisa & Netti Herawati	184-190
FUNGSI TRANSFER HUBUNGAN PERUBAHAN JUMLAH UANG BEREDAR DAN TINGKAT INFLASI Mohammad Masjkur, Yenni Anggraini dan Febrina Handayani	191-202
IDEAL RING DERET PANGKAT TERITLAK MIRING Ahmad Faisol	203-207
COMPUTATIONAL ASPECT OF WADR1 AND WADR2 ALGORITHMS FOR THE MULTI PERIOD DEGREE CONSTRAINED MINIMUM SPANNING TREE PROBLEM Wamiliana, Dwi Sakethi, and Restu Yuniarti	208-214
SEMIPARAMETRIC CANONICAL ANALYSIS METHOD FOR VARIABLES HAVE NON-LINEAR RELATIONSHIP Shiddiq Ardhi Irawan	215-229
RUANG OPERATOR $(C_{ESS}(X, L_2(a, b)), \ \cdot\ _{C_{ESS}})$ Muslim Ansori	230-234
CONVERGENCE AND CONTINUITY OF VECTORS SEQUENCES IN <i>n</i> -INNER PRODUCT SPACES Dorrah Azis - Fajar Mustaqim	235-239
EVALUASI TINGKAT KEBUTUHAN SISTEM TEMU KEMBALI INFORMASI UNIVERSITAS STUDI KASUS: UNIVERSITAS LAMPUNG Didik Kurniawan	240-248

## ANALISIS DISKRIMINAN LINIER ROBUST MENGGUNAKAN METODE VOLUME ELLIPSOID MINIMUM

Khoirin Nisa<sup>1</sup> & Netti Herawati<sup>1</sup>

<sup>1</sup> Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Universitas Lampung  
Jl. Prof. Dr. Soemantri Brojonegoro No.1 Bandar Lampung 35145  
Email : Nisa@unila.ac.id

### ABSTRACT

Discriminant Analysis is a technique concerned with separating distinct sets of objects or observations into groups and allocating new objects into previously defined groups. In the classical approach, discriminant rules are often based on the empirical mean and covariance matrix of the data. But these estimates can be sensitive to outlying data, and they become inappropriate at contaminated data sets. In this paper, we consider the problem of robust Linear Discriminant Analysis (LDA) by inserting robust estimates of mean and covariance matrix using Minimum Volume Ellipsoids (MVE) estimator. We measure the performance of classical and robust LDA through the total probability of misclassification or the Error Rate as a performance criterion. This paper shows that LDA based on MVE can improve the error rate of discriminant analysis.

*Keywords : Robust, Linear Discriminant Analysis, Minimum Volume Ellipsoids*

### PENDAHULUAN

Permasalahan dalam analisis diskriminan adalah bagaimana mengklasifikasikan objek ke dalam populasi-populasi berbeda dengan menggunakan suatu aturan pengklasifikasian yang dibangun berdasarkan data *training* yang telah diketahui asal populasinya. Analisis diskriminan linier adalah kasus di mana matriks kovarians setiap populasi diasumsikan bernilai sama. Pada analisis diskriminan linier klasik, perhitungan fungsi diskriminan linier didasarkan pada vektor nilai tengah sampel ( $\bar{\mathbf{x}}$ ) dan matriks kovarian sampel ( $\mathbf{S}$ ) biasa, yaitu  $\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_i$  sebagai elemen vektor  $\bar{\mathbf{x}}$  dan  $s_{jk} = \frac{1}{n-1} \sum_{i=1}^n (x_{ji} - \bar{x}_j)(x_{ki} - \bar{x}_k)$ ,  $j = 1, 2, \dots, p$ ,  $k = 1, 2, \dots, p$  sebagai elemen matriks  $\mathbf{S}$ . Vektor nilai tengah dan matriks kovarian ini akan optimal jika data berasal dari suatu distribusi normal multivariat, tetapi sangat sensitif terhadap adanya pencilan (*outlier*), terutama jika data mengandung pencilan yang ekstrim yang mengakibatkan distribusi data menjadi sangat menjulur (*heavy tailed distribution*), pada kasus seperti ini  $\bar{\mathbf{x}}$  dan  $\mathbf{S}$  akan kehilangan efisiensinya (Oja, 2002). Sehingga mempengaruhi fungsi diskriminan yang dibangun dan dapat mengakibatkan meningkatnya peluang kesalahan klasifikasi atau biasa disebut juga *error rate*. Untuk masalah ini diperlukan analisis diskriminan *robust* yang tegar terhadap pencilan.

Berbagai penelitian tentang analisis diskriminan *robust* telah dilakukan oleh banyak peneliti selama beberapa dasawarsa terakhir, diantaranya oleh Hawkins & McLachlan (1997), He & Fung (2000), Croux & Dehon (2001) dan Hubert & Van Driessen (2004). Pada tulisan ini kami membahas tentang analisis diskriminan linier (ADL) *robust* berdasarkan metode penduga *robust* bagi  $\bar{\mathbf{x}}$  dan  $\mathbf{S}$  yaitu metode Volume Elipsoid

Minimum (VEM) yang diperkenalkan oleh Rousseeuw pada tahun 1983 ( Hubert, Rousseeuw, Van Aelst, 2004). Penduga VEM bagi  $\bar{\mathbf{x}}$  dan  $\mathbf{S}$  diperoleh dari subhimpunan data berukuran  $h$  yang memiliki volume elipsoid terkecil.

Untuk menyederhanakan masalah, dalam tulisan ini kami menggunakan analisis diskriminan untuk dua kelompok. Misalkan terdapat himpunan  $n_1$  buah pengamatan  $X_1 = \{x_{11}, x_{12}, \dots, x_{1n_1}\}$  dalam ruang berdimensi  $p$  yang berasal dari populasi pertama  $\Phi_1 \sim N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$  dan himpunan  $n_2$  buah pengamatan  $X_2 = \{x_{21}, x_{22}, \dots, x_{2n_2}\}$  dalam ruang berdimensi  $p$  yang berasal dari populasi kedua  $\Phi_2 \sim N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$ . Maka dugaan fungsi diskriminan linier Fisher diberikan oleh:

$$L(\mathbf{x}) = \mathbf{x}^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$$

dan nilai batas: 
$$h = \frac{1}{2} (\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$$

Sebuah pengamatan baru  $\mathbf{x}$  akan diklasifikasikan ke dalam  $\Phi_1$  jika  $L(\mathbf{x}) > h$ , dan diklasifikasikan ke dalam  $\Phi_2$  jika sebaliknya. (Mangku, 2004). Dalam prakteknya fungsi diskriminan di atas dihitung berdasarkan nilai dugaan  $\hat{\boldsymbol{\mu}}$  dan  $\hat{\boldsymbol{\Sigma}}$  yang diperoleh dari data sampel *training*.

### Metode Volume Elipsoid Minimum

Penduga Volume Elipsoid Minimum (VEM) pertama kali diperkenalkan Peter J. Rousseeuw pada tahun 1983, metode ini sudah dikenal secara luas dan digunakan dalam mendeteksi pencilan pada data multivariat. Pada prinsipnya metode ini adalah mencari ellipsoid dengan volume paling minimum yang melingkupi suatu subhimpunan dari minimal  $h$  pengamatan. Subhimpunan berukuran  $h$  ini disebut *halfset* karena  $h$  sering dipilih lebih dari setengah  $n$  pengamatan. Penduga nilai tengah adalah pusat ellipsoid secara geometris dan penduga matriks kovarian adalah matriks pembentuk ellipsoid.

Mencari penduga VEM secara esensial dilakukan dengan dua proses. Bagian pertama mencari *halfset* terbaik yang memuat  $h$  pengamatan. Bagian kedua mencari volume paling minimum dari ellipsoid yang melingkupi *halfset*. Untuk sebuah *halfset* terdiri dari banyak ellipsoid yang melingkupinya. Misal terdapat data sampel berukuran  $n$  dengan  $p$  buah variabel teramati. Maka penduga VEM didefinisikan sebagai pasangan vektor nilai tengah  $\bar{\mathbf{X}}$  dan matriks varian-kovarian  $\mathbf{S}$  sedemikian sehingga determinan  $\mathbf{S}$  diminimalkan untuk memenuhi,

$$\#\{i; (\mathbf{x}_i - \bar{\mathbf{X}})' \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{X}}) \leq c^2\} \geq h$$

dengan  $h = [(n+p+1)/2]$ ,  $c^2$  adalah suatu konstanta yang dapat dipilih sebagai nilai  $\chi_{0,5(p)}^2$  dan  $\mathbf{x}$  adalah himpunan data pengamatan (Rousseeuw and Van Zomeren, 1990,1991).

Berdasarkan definisi di atas maka mencari penduga VEM sebenarnya sulit dilakukan dalam prakteknya. Misalkan untuk sampel berukuran  $n = 30$  dan dimensi data  $p = 3$ , maka  $h = (30+3+1)/2 = 17$ , sehingga terdapat  $30!/(13! \times 17!) = 119.759.850$  *halfset* yang berpotensi menjadi basis bagi penduga VEM. Karena perhitungannya yang sulit, maka Rousseeuw dan Leroy (1987) dalam Jensen *et.al* (2004) menggunakan sebuah pendekatan untuk mencari penduga VEM dengan algoritma resampling (*resampling algorithm*).

### Algoritma Resampling

Algoritma resampling didefinisikan sebagai proses mengambil sejumlah subsampel secara acak, yang memuat  $h$  pengamatan. Untuk setiap subsampel, dihitung vektor nilai tengah dan matriks kovariannya. Ellipsoid dengan volume terkecil selanjutnya digunakan untuk menduga  $\mu$  dan  $\Sigma$ .

Menurut Rousseeuw and Van Zomeren (1991), langkah-langkah dalam algoritma resampling dilakukan sebagai berikut;

1. Buat subsampel secara acak  $H = \{i_1, \dots, i_h\}$
2. Hitung nilai tengah dan matriks kovarian subsampel :

$$\bar{\mathbf{x}}_H = \frac{1}{h} \sum_H \mathbf{x}_i$$

$$\mathbf{S}_H = \frac{1}{h-1} \sum_H (\mathbf{x}_i - \bar{\mathbf{x}}_H)' (\mathbf{x}_i - \bar{\mathbf{x}}_H)$$

3. Untuk semua pengamatan, hitung jarak Mahalanobis menggunakan  $\bar{\mathbf{x}}_H$  dan  $\mathbf{S}_H$  dengan rumus:

$$D_H = \sqrt{(\mathbf{x}_i - \bar{\mathbf{x}}_H)' \mathbf{S}_H^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}_H)}$$

4. Tentukan  $\delta_H$  yaitu jarak ke- $h$  dari seluruh  $D_H$  setelah jarak diurutkan dari yang terpendek sampai jarak terjauh.
5. Hitung volume ellipsoid dengan rumus  $(V_H) = \sqrt{(\delta_H)^p \det(\mathbf{S}_H)}$
6. Ulangi langkah 1-5 dengan subsampel yang berbeda sebanyak  $m$  kali. Banyaknya  $m$  ditentukan memenuhi peluang sebagai berikut:

$$1 - (1 - (1 - \varepsilon)^{p+1})^m \geq P_0 \text{ dengan } \varepsilon = 0,5 \text{ dan } P_0 = 0,95$$

7. Simpan  $H$  dengan volume  $V_H$  paling minimum dari  $m$  pengulangan.
8. Himpunan data  $H$  dengan volume paling minimum dari  $m$  subsampel digunakan untuk menduga vektor nilai tengah dan matriks kovarian *robust* dengan rumus;

$$\bar{\mathbf{X}} = \bar{\mathbf{X}}_H$$

$$\mathbf{S} = c_{n,p} (\chi_{0.5(p)}^2)^{-1} \delta_H^2 \mathbf{S}_H$$

Dengan  $c_{n,p}$  adalah faktor koreksi yang ditentukan oleh banyaknya jumlah

variabel  $p$  dan  $n$  pengamatan, yaitu  $c_{n,p} = \left( 1 + \frac{p+1}{n-p} + \frac{1}{n-h-p} \right)^2$ .

Sedangkan  $\chi_{p,0.50}^2$  menyatakan faktor koreksi untuk menentukan konsistensi data yang diasumsikan berasal dari distribusi normal multivariat dan memiliki titik *breakdown* maksimum sebesar 50%.

## METODE PENELITIAN

Untuk mengetahui ketegaran analisis diskriminan linier robust menggunakan metode VEM kami melakukan studi secara empiris melalui simulasi data dengan bantuan perangkat lunak SAS/IML. Tahapan simulasi data yang kami lakukan adalah sebagai berikut:

1. Membangkitkan data berdistribusi normal p-variat ( $p = 2, 4$  dan  $6$ ) untuk dua kelompok yaitu  $\mathbf{X}_1^{(0)} \sim MN(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$  untuk kelompok 1 dan  $\mathbf{X}_2^{(0)} \sim MN(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$  untuk kelompok 2 masing-masing berukuran  $n = 40, 80, 120$  dan  $200$ . Semua elemen  $\boldsymbol{\mu}_1$  ditentukan bernilai 3 sedangkan semua elemen  $\boldsymbol{\mu}_2$  bernilai 10. Matriks kovarian  $\boldsymbol{\Sigma}$  dibangkitkan secara acak menggunakan SAS.
2. Membangkitkan matriks pencilan ( $\mathbf{Xout}^{(s)}$ ) yang elemennya adalah nol kecuali beberapa elemen objek yang dijadikan pencilan. Elemen pencilan dibangkitkan dari distribusi normal  $(50, (1)^2)$ . Banyaknya pencilan yang diberikan adalah 0%, 5%, 10%, 15%, 20% dan 25%.
3. Menambahkan matriks  $\mathbf{Xout}^{(s)}$  pada matriks  $\mathbf{X}_1^{(0)}$  dan  $\mathbf{X}_2^{(0)}$  sehingga diperoleh matriks  $\mathbf{X}_1$  dan  $\mathbf{X}_2$  yang sudah terkontaminasi :

$$\begin{aligned}\mathbf{X}_1^{(s)} &= \mathbf{X}_1^{(0)} + \mathbf{Xout}^{(s)} \\ \mathbf{X}_2^{(s)} &= \mathbf{X}_2^{(0)} + \mathbf{Xout}^{(s)}\end{aligned}$$

4. Menghitung vektor nilai tengah dan matriks kovarian klasik dan *robust* dari data pada  $\mathbf{X}_1^{(s)}$  dan  $\mathbf{X}_2^{(s)}$ .
5. Menduga koefisien fungsi diskriminan Fisher untuk metode klasik dan metode VEM berdasarkan vektor nilai tengah dan matriks kovarian yang diperoleh.
6. Menghitung peluang salah klasifikasi dengan metode Resubstitusi untuk hasil analisis diskriminan linier dengan metode klasik dan VEM dan DKM dengan menggunakan rumus :

$$\hat{P} = \frac{n_1}{n_1 + n_2} \hat{P}_1(R) + \frac{n_2}{n_1 + n_2} \hat{P}_2(R)$$

di mana  $\hat{P}_1(R) = \frac{1}{n_1} \sum_{j=1}^{n_1} C[1, G(x_{1j}, T)]$  dan

$\hat{P}_2(R) = \frac{1}{n_2} \sum_{j=1}^{n_2} C[2, G(x_{2j}, T)]$ , dengan  $G(x_{1j}, T)$  merupakan hasil

pengklasifikasian suatu unit dalam  $\Phi_1$ , sedangkan  $G(x_{2j}, T)$  merupakan hasil pengklasifikasian suatu unit dalam  $\Phi_2$ , dan  $C(i, j)$  adalah kriteria perhitungan dengan ketentuan sebagai berikut :

$$C(i, j) = \begin{cases} 0, & \text{jika } i = j \\ 1, & \text{jika } i \neq j \end{cases} \quad \text{untuk setiap } i \text{ dan } j$$

Selanjutnya berdasarkan hasil simulasi kami membandingkan nilai peluang kesalahan klasifikasi dari analisis diskriminan linier dengan metode klasik dan metode VEM.

### HASIL DAN PEMBAHASAN

Berdasarkan hasil simulasi dengan beberapa ukuran sampel ( $n$ ) dan dimensi data ( $p$ ) yang dicobakan diperoleh nilai peluang total kesalahan klasifikasi seperti pada Tabel 1 berikut.

**Tabel 1.** Peluang kesalahan klasifikasi metode klasik dan metode VEM pada setiap  $n$ ,  $p$  dan prosentase pencilan.

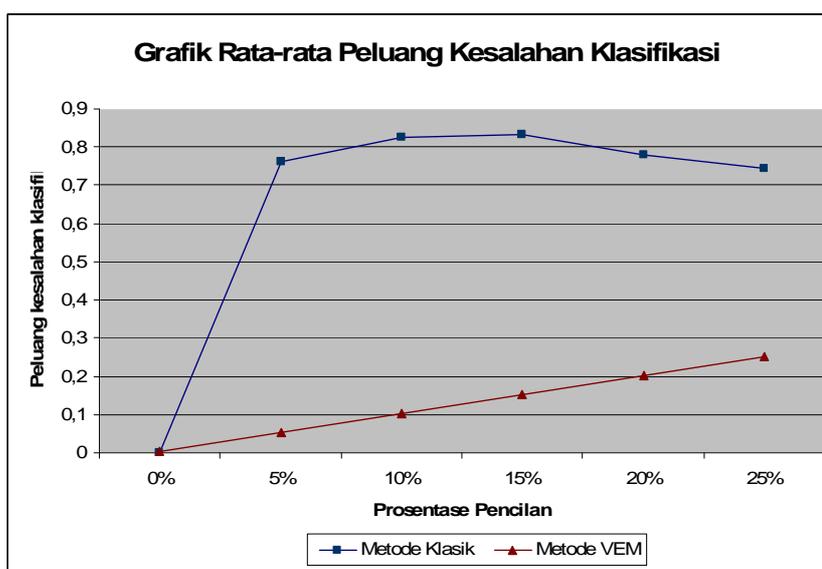
n	Prosentase Pencilan	Metode Klasik			Metode Volume Ellipsoid Minimum		
		p=2	p=4	p=6	p=2	p=4	p=6
40	0%	0	0	0	0	0	0
	5%	0,9	0,55	0,475	0,05	0,05	0,05
	10%	0,725	0,75	0,65	0,1	0,1	0,1
	15%	0,85	0,8	0,8	0,15	0,15	0,15
	20%	0,775	0,725	0,7	0,2	0,2	0,2
	25%	0,75	0,75	0,7	0,25	0,25	0,25
80	0%	0	0	0	0,0125	0	0
	5%	0,725	0,7	0,7	0,05	0,05	0,05
	10%	0,9	0,85	0,8325	0,1	0,1	0,1
	15%	0,85	0,85	0,8	0,15	0,15	0,15
	20%	0,8	0,8	0,7375	0,2	0,2	0,2
	25%	0,75	0,75	0,75	0,25	0,25	0,25
120	0%	0	0	0	0,008333	0	0
	5%	0,941667	0,883333	0,7	0,058333	0,05	0,05
	10%	0,916667	0,9	0,866667	0,1	0,1	0,1
	15%	0,85	0,85	0,85	0,15	0,15	0,15
	20%	0,8	0,8	0,8	0,2	0,2	0,2
	25%	0,75	0,75	0,75	0,25	0,25	0,25
200	0%	0,01	0	0	0,01	0	0
	5%	0,94	0,87	0,74	0,06	0,05	0,05
	10%	0,9	0,885	0,745	0,115	0,1	0,1
	15%	0,85	0,835	0,825	0,165	0,15	0,15
	20%	0,8	0,8	0,8	0,225	0,2	0,2
	25%	0,75	0,75	0,75	0,265	0,25	0,25

Berdasarkan hasil di atas terlihat bahwa peluang kesalahan klasifikasi dari analisis diskriminan dengan metode VEM jauh lebih kecil dibandingkan dengan metode klasik. Dan semakin besar jumlah pencilan dalam data akan semakin memperbesar peluang kesalahan klasifikasi metode VEM. Sedangkan analisis diskriminan dengan metode klasik tidak memiliki pola yang tetap seiring meningkatnya prosentase pencilan. Pada tabel di atas juga terlihat bahwa semakin bertambahnya variabel dalam data cenderung menurunkan nilai peluang kesalahan klasifikasi metode klasik, namun tidak demikian halnya dengan metode VEM. Analisis diskriminan linier dengan VEM terlihat menghasilkan peluang kesalahan klasifikasi yang relatif stabil untuk jumlah variabel yang berbeda.

Untuk memperjelas pengaruh jumlah pencilan terhadap peluang kesalahan klasifikasi metode VEM dan metode klasik, berikut kami sajikan rata-rata peluang salah klasifikasi metode klasik dan metode VEM dari berbagai kombinasi ukuran sample  $n$  dan jumlah variable  $p$  pada Tabel 2 dan disajikan dalam bentuk grafik batang pada Gambar 1.

**Tabel 2.** Rata-rata peluang kesalahan klasifikasi metode klasik dan VEM.

Prosentase Pencilan	Metode	
	Klasik	VEM
0%	0,000833	0,002569
5%	0,760417	0,051528
10%	0,826736	0,10125
15%	0,834167	0,15125
20%	0,778125	0,202083
25%	0,745833	0,25125



**Gambar 1.** Grafik Rata-Rata Peluang Kesalahan Klasifikasi metode Klasik dan VEM.

Pada Tabel 2 dan Gambar 1 di atas terlihat bahwa jika data tidak mengandung pencilan maka metode klasik dan metode VEM memberikan peluang kesalahan klasifikasi yang hampir sama baiknya. Namun jika terdapat pencilan pada data maka metode VEM memberikan hasil yang jauh lebih baik dari metode klasik.

### KESIMPULAN

Berdasarkan kajian yang telah kami lakukan maka dapat disimpulkan metode Volume Ellipsoid Minimum merupakan metode *robust* yang cukup baik untuk mengatasi masalah pencilan pada analisis diskriminan linier.

### DAFTAR PUSTAKA

- Croux, C. & Dehon, C. (2001) Robust linear discriminant analysis using S-estimators. *The Canadian Journal of Statistics*, 29:473–492.
- Filzmoser P. (2005). Identification of Multivariate Outliers: A Performance Study. *Austrian Journal Of Statistics*. Volume 34, No.2, hal 127–138

- Filzmoser P., Garrett R.B & Reimann C. (2005). Multivariate outlier detection in exploration geochemistry. *Computers & Geosciences*.
- Filzmoser P. (2004). A Multivariate Outlier Detection Method. In : Aivazian S., Filzmoser P. & Kharin Y. (editors). Proceedings Of The Seventh International On Computer Data Analysis and Modelling : 18-22.
- Hawkins, D. M. & McLachlan, G. J.(1997). High-breakdown linear discriminant analysis. *Journal of the American Statistical Association*, 92(437):136–143.
- He, X. & Fung, W. K. (2000) High breakdown estimation for multiple populations with applications to discriminant analysis. *Journal of Multivariate Analysis*, 72:151–162, 2000.
- Hubert, M & Van Driessen, K. (2004) Fast and robust discriminant Analysis. *Computational Statistics and Data Analysis*, 45:301–320.
- Hubert M., Rousseeuw P.J. & Van Aelst, S. (2004), Robustness. *Encyclopedia of Actuarial Sciences* Sundt B. & Teugels J. (editor), Wiley, New York, pp. 1515-1529
- Johnson, R.A. and Wichern, D.W. (1990). *Applied Multivariate Statistical Analysis*. Prentice Hall, New Jersey.
- Mangku, I.W. (2004). Estimating the Probability of Misclassification in Two-Groups Discriminant Analysis. *Jurnal Matematika dan Aplikasinya*. Vol 3, No.1.
- Oja H. (2002). Robust And Nonparametric Multivariate Methods .Department of Mathematic and Statistics University of Jyväskylä. Finland
- Olive D. J. (2004). A resistant estimator of multivariate location and dispersion. *Computational Statistical Data Analysis*. Vol 46 : 99-102.
- Rousseeuw P.J. & van Zomeren B.C. (1990). Unmasking multivariate outliers and leverage points, *Journal of the American Statistical Association*, 85:633-639.
- Rousseeuw, P.J. and van Zomeren, B.C. (1991). Robust Distance: Simulations and Cutoff Values. *The IMA Volumes in Mathematics and its Application*. 34:195- 203.