

INTISARI

ANALISIS DISKRIMINAN LINEAR *ROBUST* MENGGUNAKAN ESTIMATOR *MINIMUM COVARIANCE DETERMINANT*

Oleh

Wiwin Amalia Ritonga
13/352747/PA/15697

Analisis diskriminan adalah salah satu teknik data multivariat yang bertujuan untuk mengklasifikasikan suatu objek ke dalam suatu kelompok yang sudah ada sebelumnya. Analisis diskriminan linear mempunyai asumsi seperti variabel independen berdistribusi normal multivariat dan kesamaan matriks varians-kovarians (homokedatisitas). Apabila terdapat data *outlier*, maka analisis diskriminan klasik tidak dapat bekerja dengan baik karena akan mempengaruhi matriks varians-kovarians. Oleh sebab itu, diperlukan suatu estimator yang *robust* agar analisis diskriminan tetap optimal dalam hasil pengklasifikasiannya. Dalam skripsi ini, digunakan metode analisis diskriminan dengan menggunakan penaksir *Minimum Covariance Determinant (MCD)*. Penaksir MCD merupakan kovariansi dari sebagian pengamatan yang meminimumkan determinan matriks kovariansi. Selain itu, dalam perhitungan estimator MCD digunakan algoritma *fast-MCD* agar analisis dapat dilakukan dalam data berdimensi besar. Analisis diskriminan linear *robust* dengan menggunakan algoritma *fast-MCD* pada skripsi ini diaplikasikan untuk menentukan status seseorang terhadap ada atau tidaknya virus *dengue* dalam darah. Data diperoleh dari Rumah Sakir dr. Asmir di Salatiga pada bulan Desember 2014 s.d Januari 2015. Untuk mengukur tingkat kesalahan klasifikasi, digunakan metode *Apparent Error Rate (APER)*. Hasil analisis menunjukkan bahwa klasifikasi dengan menggunakan analisis diskriminan linear *robust* dengan menggunakan MCD memberikan hasil yang lebih baik dari pada analisis diskriminan linear klasik yang ditunjukkan oleh nilai APER yang kecil.

Kata kunci : Analisis Diskriminan Linear, Analisis Diskriminan Linear *Robust*, *outlier*, *Minimum Covariance Determinant*, *fast-MCD*, *Apparent Error Rate (APER)*

ABSTRACT

ROBUST LINEAR DISCRIMINANT ANALYSIS USING MINIMUM COVARIANCE DETERMINANT ESTIMATOR

By

Wiwin Amalia Ritonga

13/352747/PA/15697

Discriminant analysis is one of multivariate data technique to classify new observations into one of the known groups. Linear discriminant analysis has assumptions such as independent variables are multivariate distribution normal and equal variance-covariance matrices (homoscedasticity). If there are data outliers, then the classical discriminant analysis do not work well because it will affect variance-covariance matrices. Therefore required a robust estimator in order to keep optimum discriminant analysis result. In this essay is used discriminant analysis robust using minimum covariance determinant estimator. MCD estimator is covariance from some observations that minimize the determinant of variance-covariance matrices. The algorithm to calculate MCD estimators is fast-MCD because it can be use in large dimensions. This essay applies robust linear discriminant analysis using algorithm fast-MCD to classify new observations to there is a dengue virus in their blood or not. Data obtained from dr. Asmir Hospital in Salatiga in December 2014 until January 2015. Apparent Error Rate (APER) used to measure the rate of misclassification. The analysis result show that classification using robust linear discriminant analysis is better than classic linear discriminant analysis as shown by the smallest APER.

Keywords : Linear Discriminant Analysis, Robust Linear Discriminant Analysis, outlier, Minimum Covariance Determinant, fast-MCD, Apparent Error Rate (APER)