

## ABSTRAK TESIS

1. Judul Tesis : Perbandingan Metode Klasifikasi Analisis Diskriminan dan *Classification and Regression Trees (CART)*

2. Subjek : 1. CART  
                  2. Diskriminan  
                  3. Klasifikasi  
                  4. Simulasi  
                  5. Perbandingan

3. Nama : Kharis Komarudin

4. Nomor Pokok Mahasiswa : 140720090034

5. Program Studi : Statistika Terapan

6. Konsentrasi : Statistika Sosial

7. Tim Pembimbing : 1. Septiadi Padmadisatra, Ph.D  
                          2. Nusar Hajarisman, M.Si

8. Tahun Kelulusan : 2010

9. Abstrak :

Tujuan utama dari penelitian ini adalah membandingkan tingkat ketepatan klasifikasi antara metode Analisis Diskriminan dan metode *Classification and Regression Trees (CART)* dengan menggunakan data simulasi bangkitan yang ditentukan dengan dua kondisi yaitu; pertama data mengikuti distribusi normal bivariat dengan memperhatikan aspek korelasinya, kedua adalah membangkitkan data dengan kondisi tidak berdistribusi normal (eksponensial). Data simulasi dalam penelitian ini juga memperhitungkan data pencilan yang dapat menambah kesalahan klasifikasi. Adapun data pencilannya untuk masing-masing kelompok sebesar 10 persen. simulasi data yang digunakan merupakan empat tingkat ukuran sampel, yaitu  $n_1=100$ ,  $n_2=200$ ,  $n_3=500$ ,  $n_4=1000$ , semuanya akan dilakukan replikasi sebanyak 100 kali untuk melihat kekonsistensian masing-masing metode dalam pengklasifikasian.

Dari hasil penelitian menunjukkan bahwa secara umum metode CART relatif lebih baik daripada metode Analisis Diskriminan. Hal ini dapat dilihat dari persentase tingkat kesalahan klasifikasi metode CART lebih kecil jika dibandingkan dengan metode Diskriminan dan selalu konsisten pada setiap tingkat ukuran sampel.

10. Abstract :

*The main purpose of this study was to compare the level of classification accuracy between Discriminant Analysis and Classification and Regression Trees (CART) using simulated data generation defined by two conditions, namely: first the data follow a bivariate normal distribution with consideration to the correlation, the second is to generate data with the condition are not normally distributed (exponential). Simulation data in this study also calculates the outlier data that can enhance the classification error. As for outlier data for each group is 10 percent. Simulation data are also used four levels of sample size  $n_1 = 100$ ,  $n_2 = 200$ , and  $n_3 = 500$ ,  $n_4 = 1000$ , everything will be replicated as much as 100 times to see the consistency of each method in the classification.*

*The result showed that in general the CART method is relatively better than the Discriminant method. It can be seen from the percentage rate of misclassification CART method is smaller compared to the method of Discriminant and consistently at every level of the sample size.*