

A COMPARATIVE STUDY ON VARIABLE SELECTION IN MICROARRAY CLASSIFICATION

SIRIKUL LAOSRIVICHIT 5537223 EGTI/M

M.Sc. (TECHNOLOGY OF INFORMATION SYSTEM MANAGEMENT)

**THESIS ADVISORY COMMITTEE: SOTARAT THAMMABOOSADEE, Ph.D.,
SUPAPORNKIATTISIN, Ph.D., WARANYUWONGSEREE, Ph.D.**

ABSTRACT

This thesis proposes using the shrinkage method for logistic regression. The normal exponential gamma (NEG) distribution is a Bayesian-inspired method. This method uses the normal-exponential-gamma prior for coefficients of model. The typical problems of microarray data have been solved by the LASSO and the elastic net. The comparative study of the two famous methods, NEG distribution and double exponential (DE) distribution were measured by the number of selected variables, predictive accuracy, deviance and computational time. The paired sample t-test was used to analyze the mean difference between two values. The results showed that the NEG distribution was more efficient than the LASSO, the elastic net and the DE distribution.

**KEY WORDS: NORMAL EXPONENTIAL GAMMA DISTRIBUTION / LASSO /
ELASTIC NET /DOUBLE EXPONENTIAL DISTRIBUTION/ DNA
MICROARRAY**

74pages