

Understanding Mahalanobis Distance Criterion for Feature Selection

Maz Jamilah Masnan^a, Nor Idayu Mahat^b, Ali Yeon Md Shakaff^c, Abu Hassan Abdullah^c, Nur Zawatil Ishqi Zakaria^c, Nurlisa Yusuf^c, Norazian Subari^d, Ammar Zakaria^c and Abdul Hallis Abdul Aziz^c

^a Institut Matematik Kejuruteraan, Universiti Malaysia Perlis, Kampus Pauh Putra, 02600 Arau, Perlis, Malaysia

^b School of Quantitative Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia

^c Centre of Excellence for Advanced Sensor Technology, Universiti Malaysia Perlis, Taman Muhibbah, 02600 Jejawi, Perlis, Malaysia

^d Faculty of Electrical & Electronics Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia

ABSTRACT

Distance criteria are widely applied in cluster analysis and classification techniques. One of the well known and most commonly used distance criteria is the Mahalanobis distance, introduced by P. C. Mahalanobis in 1936. The functions of this distance have been extended to different problems such as detection of multivariate outliers, multivariate statistical testing, and class prediction problems. In the class prediction problems, researcher is usually burdened with problems of excessive features where useful and useless features are all drawn for classification task. Therefore, this paper tries to highlight the procedure of exploiting this criterion in selecting the best features for further classification process. Classification performance for the feature subsets of the ordered features based on the Mahalanobis distance criterion is included.

DOI: 10.1063/1.4915708