ENSEMBLE METHOD HIT RATIO FOR ROBUST TESTS OF SPREAD

T. SIN YIN[1], A. R. OTHMAN[2]
[1,2]School of Distance Education, University Science Malaysia
11800 USM Pulau Pinang, Malaysia
E-mail: [1]syin.teh@gmail.com, [2]oarahman@usm.my

ABSTRACT

An ensemble method was used to combine the outputs of several diverse classifiers to form a potentially stronger solution in ensemble system. The SAS system facilitates the building of program that can perform a composite method that combined logistic regression and discriminant analysis by using PROC LOGISTIC, PROC DISCRIM, DATA step, PROC FREQ and other SAS functions. We aim to improve the correct classification rate from the same data set. To achieve this, classification from logistic regression and discriminant analysis were integrated to form the ensemble. We took the averages of posterior probabilities (for the target values) from logistic regression and discriminant analysis, and classified according to the average posterior probabilities. Then, we created correct classification tables by defining predicted values based on the prior probability from the average posterior probabilities. The intended audiences for this paper are those who have working knowledge of Base SAS and have the fundamental grasp of statistics.

Keywords: Ensemble method, posterior probabilities, leave-one-out (L-O-O) classification, hit ratio.