SEQUENTIAL FORWARD FEATURE SELECTION WITH LOW COMPUTATIONAL COST (MonAmOR3)

✷ Author(s) :

Dimitrios Ververidis (Aristotle University of Thessaloniki, Greece)
Constantine Kotropoulos (Aristotle University of Thessaloniki, Greece)

✷ Abstract :

This paper presents a novel method to control the number of crossvalidation repetitions in sequential forward feature selection algorithms. The criterion for selecting a feature is the probability of correct classification achieved by the Bayes classifier when the class feature probability density function is modeled by a single multivariate Gaussian density. Let the probability of correct classification achieved by the Bayes classifier be a random variable. We demonstrate by experiments that the probability density function of the latter random variable can be modeled by a Gaussian density. Based on this observation, a method for reducing the computational burden in sequential forward selection algorithms is proposed. The method predicts the number of crossvalidation repetitions by employing a t–test to guarantee that a statistically significant improvement in the probability of correct classification is obtained by increasing the number of selected features. The proposed method is twice feaster than the sequential forward selection algorithm that uses a fixed number of crossvalidation repetitions and it maintains the performance of the sequential floating forward selection algorithm.

Menu