Solving classification problem using ensemble binarization classifier

ABSTRACT

Binarization strategy is broadly applied in solving various multi-class classification problems. However, the classifier model learning complexity tends to increase when expanding the number of problems into several replicas. One-Versus-All (OVA) is one of the strategies which transforming the ordinal multi-class classification problems into a series of two-class classification problems. The final output from each classifier model is combined in order to produce the final prediction. This binarization strategy has been proven as superior performance in accuracy than ordinal multi-class classifier model. However, learning model complexity (eg. Random Forest-RF ensemble decision trees) tends to increase when employing a large number of trees. Even though a large number of trees might produce a decent accuracy, generating time of the learning model is significantly longer. Hence, self-tuning tree parameter is introduced to tackle this matter. In such circumstances, a number of trees in the RF classifier are defined according to the number of class problem. In this paper, the OVA with self-tuning is evaluated based on parameter initialization in the context of RF ensemble decision tree. At the same time, the performance has also been compared with two classifier models such J48 and boosting for several well-known datasets.

Keyword: Ensemble decision tree; Bagging; Boosting; OVA; PAMAP2; WISDM.