Title: Ensemble Pruning via Individual Contribution Ordering

abstract

An ensemble is a set of learned models that make decisions collectively. Although they are usually more accurate than a single learner, existing ensemble methods often tend to construct unnecessarily large ensembles, which increases both the need for memory and computational cost. Ensemble pruning tackles this problem by only selecting a subset of ensemble members for prediction. The purpose of the selection is to reduce the memory and increase the response time with accuracy that is similar to or better than the original ensemble. An analysis of the accuracy/diversity trade-off in the context of ensemble pruning shows that classifiers that are more accurate and make more predictions in the minority group are more important for subensemble construction. Based on the gained insights, a heuristic metric for evaluating individual classifier's contribution that explicitly considers both accuracy and diversity is proposed in this paper. By incorporating ensemble members in decreasing order of their contributions, subensembles with better accuracies can be formed. Our algorithm is termed Ensemble Pruning via Individual Contribution ordering (EPIC). Experimental results on 26 UCI data sets show that subensembles formed by EPIC outperform the original ensemble, and outperforms a state-of-the-art ensemble pruning method Orientation Ordering (OO).