International Journal of Hybrid Intelligent Systems 3 (2006) 127–128
IOS Press

Guest-editorial

Special issue: Hybrid intelligent systems in ensembles

Marley M.B.R. Vellasco[a] and Fernando Gomide[b]
[a]*PUC-Rio, Brazil*
E-mail: marley@ele.puc-rio.br
[b]*Department of Computer Engineering and Automation, Faculty of Electrical and Computer Engineering, State University of Campinas, Campinas, Sao Paulo, Brazil*
E-mail: gomide@dca.fee.unicamp.br

Efforts to artificially produce intelligence via association of computing methodologies to cooperatively achieve higher performance system requirements have reached considerable maturity. In many fields, hybrid systems have succeeded to provide theoretical basis and methodological support to build real world applications. For instance, intelligent systems applications in areas such as modeling, control, optimization, classification, forecasting, information search and retrieval, image processing and recognition are common nowadays.

The Fifth International Conference on Hybrid Intelligent Systems HIS'05 was a major forum that brought state of the art information to researchers and practitioners that came worldwide. Neural networks, evolutionary computation, fuzzy systems, learning, search, support vector machines, clustering, classification, swarm intelligence, agents, artificial intelligence, to mention a few, were among the techniques used in hybrid systems suggested by authors in their papers.

This special issue collects contributions that nurture hybrid intelligent systems growth, emphasizing a balance between new developments and applications. The papers selected for this special issue focus on hybrid intelligent systems in ensembles. Committees or ensembles are receiving more attention from the computational intelligence community in order to improve the performance of pattern recognition systems. They produce a consensus decision that is potentially more accurate than individual models. This strategy is particularly useful when the available committee members are individually efficient and err in different regions of the feature space. After preliminary indications made by conference papers reviewers, anonymous referees did a second refereeing process. The result is a comprehensive and representative set of papers whose content mirror the quality of the conference.

The first paper, entitled *Evolutionary multiobjective optimization for the design of fuzzy rule-based ensemble classifiers*, by Ishibuchi and Nojima, presents a hybrid system, based on evolutionary multiobjective optimization algorithms, applied to the design of fuzzy rule-based ensemble classifiers. The idea is to attain a better performance by creating an ensemble, with high diversity, in a high-dimensional pattern classification problem. To reduce the curse of dimensionality problem, that is, the exponential increase in the total number of possible rules with the increase in the number of attributes, authors proposed a two-stage fuzzy rule selection scheme, where, in the first stage, a pre-specified number of promising fuzzy rules are generated as candidate rules, and in the second stage a genetic rule selection algorithm is applied to the set of candidate rules to construct fuzzy rule-based classifiers with high accuracy and high interpretability. The proposed hybrid system was applied to six high-dimensional data sets, providing interesting results.

In the second paper – *Using weighted dynamic classifier selection methods in ensembles with different levels of diversity*, Canuto et al. investigate the use of confidence measures (weights) in selection-based methods applied to multi-classifier systems. The aim is to smooth out the dependency of these methods to the diversity of the ensemble members. The paper evaluates the benefits of applying weights to three selection-based methods: Dynamic classifier selection; Dynamic Classifier Selection based on multiple classifier behavior; and Dynamic classifier selection using also Decision Templates. Authors also examine which confidence measure, between two proposals, is more suitable to be used with the selection-based methods. Experiments of the proposed confidence measures applied to the selection-based methods were carried out with homogeneous and heterogeneous ensemble configurations and the results obtained, with two different data sets, indicate that the use of weights improves accuracy of selection-based methods when the diversity of the ensemble is reduced.

The third paper, by Bernardini et.al., entitled *Constructing Ensembles of Symbolic Classifiers*, proposes a methodology that overcomes the lack of transparency of common ensemble combination methods. As mentioned by the authors, most well known ensemble combining methods do not provide any explanation about the ensemble's final decision. The proposed approach is based on the combination of symbolic classifiers that are evaluated using three different metrics: accuracy, Laplace accuracy and negative confidence. Authors propose three different voting mechanisms, namely: Unweighted Voting; Weighted by Mean Voting; and Weighted by Mean and Standard Error Voting. The last two methods favor hypotheses with either lower mean error or lower mean and standard errors, respectively. This paper also presents an interesting computational environment, called Ensemble Learning Environment, which includes the three proposed methods for combining classifiers, as well as methods for classifying examples. The interest point is that it can be easily extended to support other methods. The proposed methodology is evaluated with three benchmark databases (Nursery, Chess-Kr-Vs-Kp and Splice from UCI Repository), Authors demonstrate that meaningful results can be achieved using a small number of base-classifiers.

In the last paper – *A boosting approach to remove class label noise*, Karmaker and Kwek propose a modification of AdaBoost ensemble creation method that is more tolerant to class label noise. The basic idea is to enhance the prediction accuracy by reducing susceptibility to overfitting the training instances corrupted by class label noise. The proposed method, called OR-Boost, employs a threshold scheme to recognize and eliminate outliers in the data set, thus improving the final performance. The proposed methodology was applied to a wide range of synthetic as well as benchmark data sets, demonstrating that a careful selection of threshold level improves the outlier detection performance, hence increasing the final classification accuracy.

We believe that these four papers, presenting different hybrid approaches to improve ensemble classification performance, give a glance of the state of the art in ensemble methodology.

Rio de Janeiro, September 2006