



Editorial

Editorial: New trends and applications on hybrid artificial intelligence systems

This Special Issue is an outgrowth of the HAIS'10, the 5th International Conference on Hybrid Artificial Intelligence Systems, which was held in San Sebastián, Spain, 23–25 June 2010. The HAIS conference series is devoted to the presentation of innovative techniques involving the hybridization of emerging and active topics in data mining and decision support systems, information fusion, evolutionary computation, visualization techniques, ensemble models, intelligent agent-based systems (complex systems), cognitive and reactive distributed AI systems, case base reasoning, nature-inspired smart hybrid systems, bio- and neuro-informatics and their wide range of applications. It is dedicated to promote novel and advanced hybrid techniques as well as interdisciplinary applications and practice. HAIS'10 received over 269 submissions worldwide. After careful peer-review, only 132 papers were accepted for presentation at the conference and for inclusion in the proceedings, published as Springer's Lecture Notes in Artificial Intelligence series. Authors of the most innovative papers within the scope of the NEUROCOMPUTING Journal were invited to submit their substantially extended and updated papers with additional original materials based on their most recent research findings. Each submitted paper was subsequently reviewed by 3–5 experts and leading researchers in the field. Finally eighteen papers passed the journal's rigorous review process and were included in this Special Issue. They present an exclusive sample of the conference and its recent topics.

In the area of artificial vision and image processing, Segovia et al. present a comparison between two methods for analyzing PET data in order to develop more accurate CAD systems for the diagnosis of Alzheimer's disease. One of them is based on the Gaussian Mixture Model (GMM) and models the Regions Of Interest (ROIs) defined as differences between controls and AD subject. After GMM estimation using the EM algorithm, feature vectors are extracted for each image depending on the positions of the resulting Gaussians. The other method under study computes score vectors through a Partial Least Squares (PLS) algorithm based estimation and those vectors are used as features. Before extracting the score vectors, a binary mask based dimensional reduction of the input space is performed in order to remove low-intensity voxels. The validity of both methods is tested on the ADNI database by implementing several CAD systems with linear and nonlinear classifiers and comparing them with previous approaches such as VAF and PCA.

The contribution of Chyzyk et al. entitled "Hybrid Dendritic Computing with Kernel-LICA applied to Alzheimer's disease detection in MRI" presents the issue of enhancing the generalization

classification power of the Dendritic Computing approach to classifier training. The paper contributes a hybrid of the Lattice Independent Component Analysis (LICA) and the Kernel transformation that provides an enhanced generalization on cross-validation experiments performed on a dataset of magnetic resonance imaging (MRI) features for Alzheimer's disease computer aided diagnosis design.

The contribution by Cilla et al. presents a human action recognition system combining multiple camera observations. The proposal is centered on how to efficiently combine the observations from different viewpoints without the explicit usage of 3D models or camera calibration. To achieve this goal, a local action prediction is defined using the features extracted at each camera. These local predictions take the form of posterior probability distributions to capture the uncertainty on the classification. These distributions are combined to obtain a final posterior distribution on the performed action. Experiments show that the proposed scheme achieves successful results in the human recognition task.

In bioinformatics and medical applications, the contribution by López et al. presents the theoretical and practical results of a novel data mining process that combines hybrid techniques of association analysis and classical sequentiation algorithms of genomics to generate grammatical structures of a specific language. The authors used an application of a compilers generator system that allows the development of a practical application within the area of grammarware, where the concepts of the language analysis are applied to other disciplines, such as bioinformatic. The tool allows the complexity of the obtained grammar to be measured automatically from textual data. A technique of incremental discovery of sequential patterns is presented to obtain simplified production rules, and compacted with bioinformatics criteria to make up a grammar.

The work by Cuadra et al. entitled "Response Calibration in Neuroblastoma Cultures over Multielectrode Array" assesses the statistical relevance of neuroblastoma culture responses to excitation when they are grown in a Multielectrode Array (MEA) setup in order to provide an appropriate calibration of the systems response to excitation. The MAE are characterized by very low signal-to-noise ratio, which is improved by the proposed calibration, opening the possibility to employ neuroblastoma cultures for information processing in hybrid systems.

In Evolutionary Computation, Maravall et al. present the application of evolutionary strategies to the self-emergence of a common lexicon in a population of agents. By modeling the vocabulary or lexicon of each agent as an association matrix or

look-up table that maps the meanings (i.e. the objects encountered by the agents or the states of the environment itself) into symbols or signals it is checked whether it is possible for the population to converge in an autonomous, decentralized way to a common lexicon, so that the communication efficiency of the entire population is optimal.

In the contribution by García-Gutiérrez et al. a novel hybrid classifier applied to remote sensing data fusion is presented. The classifier, called EVOR-STACK, is based on the use of ensemble techniques and evolutionary computation. The latter allows to calculate a set of weights for every feature depending on the candidate label and therefore, extending the classical evolutive feature weighting concept. The former improves the final results due to the introduction of contextual data previously weighted with the results of the label-dependent evolutionary algorithm. The results confirmed that EVOR-STACK outperforms SVM and R-STACK (its predecessor, presented in HAIS-2010) when they are applied to LIDAR and images data fusion.

In terms of Neural-based models and applications, Fernández-Navarro et al. propose an alternative to the standard Gaussian Radial Basis Function for binary classification problems. The authors present q-Gaussian Radial Basis Function Neural Networks, where the basis functions include a supplementary degree of freedom in order to adapt the model to the distribution of the data. A Hybrid Algorithm (HA) is used to search for a suitable architecture and parameters for these models.

In the contribution by Lei and Ghorbani, two new clustering algorithms are presented, the Improved Competitive Learning Network (ICLN) and the Supervised Improved Competitive Learning Network (SICLN), for the applications in the area of fraud detection and network intrusion detection. The ICLN is an unsupervised clustering algorithm applying new rules to the Standard Competitive Learning Neural Network (SCLN). In the ICLN, network neurons are trained to represent the center of the data by a new reward-punishment update rule. The authors claim that the new update rule overcomes the instability of the SCLN. The SICLN is a supervised clustering algorithm further developed from the ICLN by introducing supervised mechanism. According to the authors, in the SICLN, the new supervised update rule utilizes data labels to guide the training process in order to achieve a better clustering result, and can be applied to both labeled and unlabeled data. They claim that the SICLN algorithm is highly tolerant to missing or delayed labels. Furthermore, the SICLN is completely independent from the initial number of clusters because it is able to reconstruct itself according to the labels of the cluster members. The authors successfully demonstrate the SICLN's high performance through experimental analysis using both academic research data and practical real-world data for fraud detection and network intrusion detection. They also show that the SICLN outperforms traditional unsupervised clustering algorithms.

Then Prieto et al. present a hybrid intelligent system to provide autonomous robots with the ability to classify the motion behavior patterns of a group of robots present in their surroundings. It is a first step in the development of a cognitive model that can detect and understand the events occurring in the environment, events that are not the robot's own actions. The hybrid system is called ANPAC (Automatic Neural-based Pattern Classifier). It uses a variable size ANN to perform pattern classification and an advisor module to adjust the preprocessing parameters and, consequently, the size of the ANN, depending on the learning results of the network. The components and operations of ANPAC are described in depth and illustrated using an example related to the recognition of behavior patterns in the flocks motion.

The next contribution by Heras et al. presents an abstract argumentation framework for the support of agreement processes

in agent societies. It takes into account some arguments, among them attacks arguments, and the social context of the agents that put forward arguments. The framework is implemented as a neural network that computes the set of acceptable arguments and can be tuned to give more or less importance to argument attacks. The proposal is illustrated with an example in a real domain of a water-rights transfer market.

Two contributions based on ensembles techniques are included in this special issue. The contribution by Buza et al. proposes a model-selection framework for stacking regressors. Due to the differences in their underlying assumptions, principles and settings, various regression models (such as neural and Bayesian networks, support vector machines, regression trees, etc.) err differently on instances, i.e., compared to the true value of the target, some of the models predict lower values, while other models predict higher ones. Therefore, different models are potentially able to compensate for each other's errors. This idea of error compensation is explored and exploited for making stacking-based ensembles more accurate: the proposed framework focuses on the principled selection of a well-performing set of models, which may be substantially different from the individually best models due to the inter-correlations between them.

Then Corchado and Baroque present a novel model for multi-dimensional data visualization in 2 dimensions. It can be considered an extension of a previously developed Visualization Induced Self-Organizing Map by the use of the ensemble meta-algorithm in an unsupervised context. The contribution shows, with a comparative study, how the calculation of slightly different ViSOM maps and its subsequent fusion into a final map can give as a result a more truthful visual 2D display of the analyzed data. This algorithm proves to be a useful tool for data visualization where the dataset under analysis is especially complex, as the enhanced representation compensates for the extra complexity of its calculations.

Regarding issues of Classification methods & Information Processing, Wilk and Wozniak present a possibility of generalizing the two-class classification into multiclass classification by means of a fuzzy inference system. Fuzzy combiner harnesses the support values from classifiers to provide final response having no other restrictions on their structure. Authors compare the combination methods with ECOC and two variations of decision templates, based on Euclidean and symmetric distance. The quality of the proposed method was evaluated via computer experiment.

Then, in the contribution by Chmielnicki and Stapor, the authors present a method of combining a support vector machine (SVM)—discriminative classifier with regularized discriminant analysis (RDA)—generative classifier. The hybrid SVM–RDA classifier is used in the protein fold prediction. It is a very challenging multiclass problem with high data dimensionality and a very small number of samples for each class. The authors show how to deal with these difficulties using advantages of generative and discriminative classifiers.

In the contribution by Kazienko and Kajdanowicz entitled “Label-dependent Node Classification in the Network”, are presented an original extension and application of sampling algorithms to node classification in the networked data. In their new approach, they make use of new input variables, calculated based on structural network measures. Additionally, these measures, called label-dependent features, are computed separately for sub-networks of nodes with a given label-class. As a result, two new approaches of sampling algorithms—LDBootstrapping and LDGibbs have been developed for the purpose of collective classification. According to experimental studies carried out, the novel approaches provide more accurate results comparing to competitive ones in generalization for sparse networked datasets.

In the contribution by Zafra et al., is presented a filter-based feature selection method for working in the multiple-instance learning scenario called ReliefF-MI. This method based on the principles of the well-known ReliefF algorithm is applied as a preprocessing step that is completely independent from the multi-instance classifier learning process and therefore is more efficient and generic than wrapper approaches proposed in this area. Different extensions are designed and implemented and their performance checked in multiple-instance learning. Experimental results on five benchmark real-world datasets and seventeen classification algorithms confirm the utility and efficiency of this method, both statistically and from the point of view of the execution time.

Finally Villar et al. propose a novel approach to represent the uncertainty in the data in order to learn white box models. This representation includes; introducing fuzzy evaluation of the imprecise variables, a genetic programming approach and the learning algorithms to deal with the data. Two main multi-objective algorithms are used: Multi-objective Simulated Annealing and NSGA-II. The obtained results show this approach as a promising path for developing new controller design techniques.

We would like to thank our peer-reviewers for their diligent work and efficient efforts. We are also grateful to the Editor-in-Chief, Prof. Tom Heskes, for his continued support for the HAIS conference and for this Special Issue on this prestigious journal.

Guest-editors

Emilio Corchado*

University of Salamanca, Salamanca, Spain

*Visiting professor at VSB-Technical University of Ostrava,
Czech Republic*

E-mail address: escorchado@usal.es

Manuel Graña

University of the Basque Country, San Sebastián, Spain

E-mail address: manuel.grana@ehu.es

Michał Woźniak

Wroclaw University of Technology, Wroclaw, Poland

E-mail address: michal.wozniak@pwr.wroc.pl

* Corresponding author.