

Contemporary training methodologies based on evolutionary artificial neural networks with product and sigmoid neurons for classification

Antonio J. Tallón-Ballesteros

Department of Languages and Computer Systems, University of Seville, Spain.

E-mail: atallon@us.es

Abstract. This thesis introduces three contributions to train feed-forward neural network models based on evolutionary computation for a classification task. The new methodologies have been evaluated in three-layered neural models, including one input, one hidden and one output layer. Particularly, two kind of neurons such as product and sigmoidal units have been considered in an independent fashion for the hidden layer. Experiments have been carried out in a good number of problems, including three complex real-world problems, and the overall assessment of the new algorithms is very outstanding. Statistical tests shed light on that significant improvements were achieved. The applicability of the proposals is wide in the sense that can be extended to any kind of hidden neuron, either to other kind of problems like regression or even optimization with special emphasis in the two first approaches.

Keywords: Artificial neural networks, evolutionary algorithm, classification, product unit, sigmoid unit, feature selection

1. Introduction

Learning algorithms can be grouped into two categories: a) black-box methods, such as neural networks or Bayesian classifiers and b) knowledge-oriented methods, such as the models created by decision trees, association or decision rules. Data Mining techniques are sensitive to the information quality on that the knowledge discovery will be carried out. The higher data quality, the higher quality decision-making models.

The goal of this paper is to get more accurate neural networks models with a greater efficiency that in previous evolutionary approaches.

2. Contributions

The three contributions of the thesis are detailed in the next subsections. The training of the neural network (NN) models is performed by means of a baseline evolutionary algorithm (EA) that simultaneously evolves the weights and the architecture of the NN. The interested reader can find the EA in Section 2.2. of [3], including the pseudo-code, their foundations and related literature. Initially, the EA was proposed in the context of product-unit neural network (PUNN) models. Taking as starting point the aforementioned EA, some novelty methodologies are presented for the training of neural networks with special attention to classification models containing product or sigmoidal neurons in the hidden layer. A deep overview of them can be read in [1] along with all the results and statistical tests.

2.1. Experimental Design Distribution(EDD)

In this first proposal [2], some parameters regarding either the topology of the PUNN model or the EA are distributed throughout all the processing elements of the computation system. More concretely, these parameters are the maximum number of neurons in the hidden layer, the maximum number of generations and the parameter value associated with the parametrical mutation. Recently, this contribution has been extended to consider sigmoidal neurons in the hidden layer and the resultant model has been named Experimental Design Distribution with Sigmoidal units (EDDSig) [4].

2.2. Two-Stage Evolutionary Algorithm (TSEA)

The second contribution diversifies the neural network architecture and is composed of two stages. During the first stage, two populations with different properties are created and evolved for a small number of generations. Next, the half best of each population are merged in a new population. In the second stage, the new population is evolved for a full evolutionary cycle. More details about this approach are provided in [3]. Also, the algorithm is available upon request. The natural extension to sigmoidal units has been called Two-Stage Evolutionary Algorithm for neural networks with Sigmoidal units (TSEASig) [4].

2.3. Two-Stage Evolutionary Algorithm with Feature Selection (TSEAFS)

This third proposal [5] involves a data preprocessing –applying feature selection methods implemented as filters– on the data set in order to improve the efficacy and the simplicity of the obtained models. Some feature selectors are independently applied to the training set of the problem at hand getting a list of attributes, that will be used to obtain the reduced training and test sets for the learning of the PUNN models.

3. Results and final remarks

The experimentation was conducted using a great deal of classification problems (refer to Appendix A). EDD was evaluated in 25 data sets and got the best averaged results followed by C4.5 and PART classifiers. TSEA was assessed in 30 problems and reached the highest test accuracy average; MLP got the second best ones. TSEAFS was tested with data sets that report a test error rate around 20% yielding a total of 18 problems; the global best test results were obtained with TSEAFS and depending on the filter the second best option was MLP or RBF.

Comparing our models utilising product or sigmoid units the achieved conclusions are as follows. First, EDD is more accurate than EDDSig, however the latter is faster than the former. Second, TSEA has a significantly better efficacy and is about a forty percent faster than EDD. Third, TSEASig has significant greater accuracy than EDDSig. Fourth, TSEAFS obtains simpler models with a dimensionality reduction greater than a fifty percent compared with TSEA.

Acknowledgements

This work has been partially subsidized by TIN2007-68084-C02-02, TIN2008-06681-C06-03 and TIN2011-28956-C02-02 projects of the Spanish Inter-Ministerial Commission of Science and Technology (MICYT), FEDER funds, P08-TIC-3745 and P11-TIC-7528 projects of the “Junta de Andalucía” (Spain).

References

- [1] A.J. Tallón-Ballesteros. *Nuevos modelos de Redes Neuronales Evolutivas para Clasificación. Aplicación a Unidades Producto y Unidades Sigmoide*. PhD thesis, University of Seville (Spain), 2013. Available at: <http://fondosdigitales.us.es/tesis/tesis/1990/nuevos-modelos-de-redes-neuronales-evolutivas-para-clasificacion-aplicacion-unidades-producto-y-unidades-sigmoide/>.
- [2] A.J. Tallón-Ballesteros, P.A. Gutiérrez-Peña, and C. Hervás-Martínez. Distribution of the search of evolutionary product unit neural networks for classification. *arXiv preprint arXiv:1205.3336*, 2012.
- [3] A.J. Tallón-Ballesteros and C. Hervás-Martínez. A two-stage algorithm in evolutionary product unit neural networks for classification. *Expert Systems with Applications*, 38(1):743–754, 2011.
- [4] A.J. Tallón-Ballesteros, C. Hervás-Martínez, and P.A. Gutiérrez. An extended approach of a two-stage evolutionary algorithm in artificial neural networks for multiclassification tasks. In *Innovations in Intelligent Machines-3*, pages 139–153. Springer, 2013.
- [5] A.J. Tallón-Ballesteros, C. Hervás-Martínez, J.C. Riquelme, and R. Ruiz. Feature selection to enhance a two-stage evolutionary algorithm in product unit neural networks for complex classification problems. *Neurocomputing*, 114:107–117, 2013.

Appendix

A. Classification problems

Appendicitis, Australian credit approval, Balance, Breast (Cancer, Tissue and Wisconsin), Cardiotocography, Heart (Statlog and disease Cleveland), Hepatitis, Horse colic, Thyroid disease (Hypothyroid and Newthyroid), Ionosphere, Labor Relations, Led24, Liver disorders, Lymphography, Parkinsons, Pima Indians diabetes, Steel Plates Faults, Molecular Biology, SPECTF, Vowel, Waveform, Wine Quality, Yeast, BTX, Listeria monocytogenes and Liver-transplantation data sets were used in the experiments.