



Master in Artificial Intelligence (UPC-URV-UB)

Master's Thesis Proposal

General Information

Master's Thesis Title: *Learning to forget. Towards a fully adaptive paradigm in ensemble learning*

Publication Date: 04/11/10

Expiry Date: 04/11/11

Modality: **research work**

M.Sc. Th. Advisor: Dr. Oriol Pujol Vila

M.Sc. Th. Advisor's Dept. & University: Matemàtica Aplicada i Anàlisi, Universitat de Barcelona

M.Sc. Th. Advisor e-mail: oriol_pujol@ub.edu

Observations:

Student's Name:
(if already known)

M.Sc. Thesis Description

Brief Description: (a 4-5 lines long paragraph)

In adaptive learning a machine learning strategy must deal with nonstationary environments while keeping the complexity of the model/solution fixed. For instance, consider the case of online learning, one needs the system to learn as data arrives, thus the system must adapt and evolve, learning from new data and forgetting part of the model if it is not useful any more. In this thesis we plan to analyze and propose a solution to this problem from the ensemble learning point of view.

Detailed Description: (a half-page or one-page description, may include a brief task planning)

Machine Learning has evolved in many different lines of research to cope with many of nowadays learning problems. However, there still are many open problems and unstudied areas of research. This is the case of adaptive learning. In adaptive learning a machine learning strategy must deal with nonstationary environments while keeping the complexity of the model/solution fixed. As mentioned formerly, consider the case of online learning, one needs the system to learn as data arrives, thus the system must adapt and evolve, learning from new data and forgetting part of the model if it is not useful any more. However, most of nowadays online machine learning models do not deal specifically with the forgetting problem. In this thesis we plan to analyze and propose a solution to this problem from the ensemble learning point of view. Ensemble learning has become one of the most fruitful and successful learning paradigms in machine learning. Algorithms such as Adaboost, Bagging, Error correcting output codes, model averaging, etc, are part of this learning branch. The main hypothesis in ensemble learning is that a set of hypotheses may be combined reducing the overall bias and variance of the result. In this work we want to focus on boosting techniques. This master thesis is oriented to those of you who want to follow the master with a PhD thesis.

Planning:

- (A) Analysis of adaptive learning strategies (i.e. Learn++)
- (B) Formulation of the new proposal from the point of view of ensemble learning.
- (C) Analysis of convergence of the method.
- (D) Validation in non-stationary and changing environments.

References: (bibliographical references, relevant web links, etc.)

L. Kuncheva, "Classifier Ensembles for Changing Environments", LNCS 3077, pp. 1-15, 2004.

R. Elwell and R. Polikar, "Incremental Learning of Variable Rate Concept Drift", MultipleClassifiers Systems 2009, pp. 142-151, 2009.

G. Widmer and M. Kubat, "Learning in presence of Concept Drift and Hidden Contexts", Machine Learning, vol. 23, pp. 69-101, 1996.

Location and Date: Barcelona,

To the Academic Commission of the Master in Artificial Intelligence (CAIMIA)