

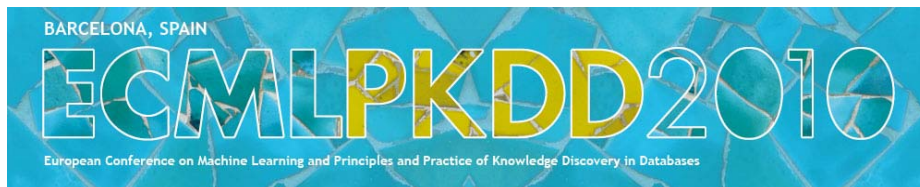
Proceedings of the First International Workshop on

Handling Concept Drift
in Adaptive Information Systems:
Importance, Challenges and Solutions

HaCDAIS 2010

Held in conjunction with
ECML/PKDD 2010

September 24, 2010
Barcelona, Spain



Editors

Mykola Pechenizkiy

Eindhoven University of Technology, the Netherlands

Indrė Žliobaitė

Eindhoven University of Technology, the Netherlands

Preface

In the real world data is often non stationary. In predictive analytics, machine learning and data mining the phenomenon of unexpected change in underlying data over time is known as concept drift. Changes in underlying data might occur due to changing personal interests, changes in population, adversary activities or they can be attributed to a complex nature of the environment.

When there is a shift in data, the predictions might become less accurate as the time passes or opportunities to improve the accuracy might be missed. Thus the learning models need to be adaptive to the changes.

The problem of concept drift is of increasing importance to machine learning and data mining as more and more data is organized in the form of data streams rather than static databases, and it is rather unusual that concepts and data distributions stay stable over a long period of time. It is not surprising that the problem of concept drift has been studied in several research communities including but not limited to machine learning and data mining, data streams, information retrieval, and recommender systems. Different approaches for detecting and handling concept drift have been proposed in the literature, and many of them have already proved their potential in a wide range of application domains, e.g. fraud detection, adaptive system control, user modeling, information retrieval, text mining, biomedicine.

HaCDAIS 2010 workshop¹ organized in conjunction with ECML/PKDD 2010 and held on 24 September 2010 in Barcelona, Spain provides a focused international forum for researchers to discuss new, we aim to attract researchers with an interest in handling concept drift and recurring contexts in adaptive information systems.

Topics discussed at the workshop include classification and clustering on data streams and evolving data, change and novelty detection in online, semi-online and offline settings, adaptive ensembles, adaptive sampling and instance selection, incremental learning and model adaptivity, delayed labeling in data streams, dynamic feature selection, handling local and complex concept drift, qualitative and quantitative evaluation of concept drift handling performance, reoccurring contexts and context-aware approaches, application-specific and domain driven approaches within the areas of information retrieval, recommender systems, pattern recognition, user modeling, decision support and adaptive information systems

These proceedings include abstract of the invited talk, invited software report and six peer-reviewed papers accepted to the workshop, two as full papers and four as short papers. We would like to thank all the authors for their interest in the workshop and for submitting their contributions.

We would like to thank PC member for publicizing the workshop and providing timely and constructive reviews for the submitted papers.

We are thankful to Prof. João Gama who kindly agreed to give an invited talk “Reasoning about the Learning Process” and Dr. Albert Bifet for presenting the “MOA: Massive Online Analysis” framework and software demonstration.

¹ For further information about workshop-related information, scientific program overview and link to the online proceedings please refer to <http://www.wis.win.tue.nl/hacdais2010/>

We are looking forward for an interesting workshop and to meeting you in Barcelona!

27 July 2010
Eindhoven, the Netherlands

Mykola Pechenizkiy
Indrė Žliobaitė

Co-Chairs of
HaCDAIS 2010 Workshop

Workshop Organization

Workshop Chairs

Mykola Pechenizkiy Eindhoven University of Technology, the Netherlands
Indrė Žliobaitė Eindhoven University of Technology, the Netherlands

Workshop Program Committee

Albert Bifet University of Waikato, New Zealand
Sarah Jane Delany Digital Media Centre, Ireland
Anton Dries Katholieke Universiteit Leuven, Belgium
Bogdan Gabrys Bournemouth University, UK
João Gama University of Porto, Portugal
Ioannis Katakis Aristotle University of Thessaloniki, Greece
Yehuda Koren Yahoo! Research, Israel
Ludmila Kuncheva Bangor University, UK
Matthijs van Leeuwen Universiteit Utrecht, the Netherlands
Ernestina Menasalvas Universidad Politecnica de Madrid, Spain
Robi Polikar Rowan University, USA
Myra Spiliopoulou Otto-von-Guericke-University Magdeburg, Germany
Alexey Tsymbal Siemens AG, Germany
Athena Vakali Aristotle University, Thessaloniki, Greece

External Reviewers

Elena Ikonomovska Jozef Stefan Institute, Slovenia

Table of Contents

Invited Talk

Reasoning about the Learning Process	1
<i>João Gama</i>	

Invited Software Report

MOA: Massive Online Analysis, a framework for stream classification and clustering	3
<i>Albert Bifet, Geo Holmes, Bernhard Pfahringer, Philipp Kranen, Hardy Kremer, Timm Jansen, and Thomas Seidl</i>	

Full Papers

Incremental option trees for handling gradual concept drift	17
<i>Elena Ikonomovska, João Gama, and Sašo Džeroski</i>	
Handling concept drift in preference learning for interactive decision making.....	29
<i>Paolo Campigotto, Andrea Passerini, and Roberto Battiti</i>	

Short Papers

Cost-sensitive boosting for concept drift.....	41
<i>Ashok Venkatesan, Narayanan C. Krishnan, Sethuraman Panchanathan</i>	
An adaptive hybrid recommender system that learns domain dynamics.....	49
<i>Fatih Aksel and Ayşenur Birtürk</i>	
Modeling the example life-cycle in an online classification learner.....	57
<i>Gary R. Marrs, Ray J. Hickey, and Michaela M. Black</i>	
An incremental text segmentation by clustering cohesion.....	65
<i>Raúl Abella Pérez and José Eladio Medina Pagola</i>	