Report of the First International Workshop on Learning over Multiple Contexts (LMCE 2014)

Cèsar Ferri
Univ. Politèc. de València
Cami de Vera, S/N 46022
València, Spain

cferri@dsic.upv.es

Peter Flach
University of Bristol
Woodland Road
Bristol BS8 1UB, U.K.

peter.flach@bristol.ac.uk

Nicolas Lachiche
ICube, Univ. Strasbourg
300 bd Brant
67400 Illkirch, France

nicolas.lachiche@unistra.fr

ABSTRACT

The first international workshop on Learning over Multiple Contexts, devoted to generalization and reuse of machine learning models over multiple contexts, was held on September 19th, 2014, as part of the 7th European machine learning and data mining conference (ECML-PKDD 2014) in Nancy, France. This short report summarizes the presentations and discussions held during the LMCE 2014 workshop, as well as the workshop conclusions and the future agenda.

1. INTRODUCTION AND MOTIVATION

Adaptive reuse of learnt knowledge is of critical importance in the majority of knowledge-intensive application areas, particularly when the context in which the learnt model operates can be expected to vary from training to deployment. In data mining, the extracted knowledge will generally be used in a new deployment context with different characteristics. Additionally the model might need adaptation, for example if there are manifestations of data shift or concept drift. This has been studied also in machine learning, for example, in relation to variations in class and cost skew in (binary) classification, leading to the development of tools such as ROC analysis for adjusting decision thresholds to operating conditions concerning class and cost skew. More recently, considerable effort has been devoted to research on transfer learning, domain adaptation, and related approaches.

Given that the main business of predictive models is to generalise from the training context to the deployment context, there is clearly scope for developing a general notion of *operating context*. Without such a notion, a model predicting 'sales in Prague for this week' may perform poorly in Nancy for next Wednesday. The operating context has changed in terms of location as well as resolution. While a given predictive model may be sufficient and highly specialised for one particular operating context, it may not perform well in other contexts. If sufficient training data for the new context is available it might be feasible to retrain a new model; however, this is generally not a good use of resources, and one would expect it to be more cost-effective to learn one general, *versatile* model that effectively generalizes over multiple contexts.

The aim of this workshop was to bring together people working in areas related to versatile models and model reuse over multiple contexts. Given the advances made in recent years on specific approaches such as transfer learning, an attempt to start developing an overarching theory is now feasible and timely, and could be expected to generate considerable interest from the machine learning and data mining communities. Papers were solicited in all areas relating to model reuse and model generalization including the following areas:

- Transfer learning
- Data shift and concept drift
- Domain adaptation
- Transductive learning
- Multi-task learning
- ROC analysis and cost-sensitive learning
- Background knowledge, relational learning
- Context-aware applications
- Incomplete information, abduction
- Meta-learning

The program committee was selected to cover all those areas, with 18 members from 8 countries: Brazil, China, France, Japan, Portugal, Spain, UK and USA.[1]

The workshop was held at the European machine learning and data mining conference in order to gather specialists from all relevant disciplines. We received submissions from 10 countries: Austria, Bangladesh, Belgium, Brazil, France, Germany, The Netherlands, Romania, Spain and UK. Each paper was reviewed by at least 2 program committee members. 13 regular papers and 3 work-in-progress papers were accepted. Around 40 people participated in the day-long workshop.

---

[1]Chowdhury Farhan Ahmed, Univ Strasbourg, France
Charles Elkan, Univ California - San Diego, USA
Cèsar Ferri, Univ Politècnica de València, Spain
Peter Flach, Univ Bristol, UK
Amaury Habrard, Univ Jean Monnet, Saint-Etienne, France
Francisco Herrera, Univ Granada, Spain
Meelis Kull, Univ Bristol, UK
Nicolas Lachiche, Univ Strasbourg, France
Dragos Margineantu, Boeing Research, USA
Adolfo Martínez-Usó, Univ Politècnica de València, Spain
Weike Pan, Shenzhen U, China
Ricardo Prudencio, Univ Pernambuco, Brazil
Joaquin Quiñonero, Facebook, USA
María José Ramírez-Quintana, Univ Politècnica de València, Spain
Carlos Soares, Univ Porto, Portugal
Masashi Sugiyama, Tokyo Institute of Technology, Japan
Bianca Zadrozny, Univ Fluminense, Brazil
Huimin Zhao, Univ Wisconsin - Milwaukee, USA

2. ORGANISATION AND SESSIONS

The technical program of LMCE 2014 consisted of two invited talks and five sessions devoted to a total of 16 contributed papers.

2.1 Invited talks

The workshop started with an invited talk by José Hernández-Orallo (Universitat Politècnica de València, Spain) on *Context Change and Versatile Models in Machine Learning*. Noting that models are often deployed in a different context than training, the speaker listed many types of contexts and proposed an explicitation of contexts with the corresponding adaptation procedures. His presentation emphasized the importance and usefulness of versatile models that can be adapted to different contexts, in an approach called *reframing*, with several examples.

Antonio M. Lopez (Universitat Autònoma de Barcelona, Spain) gave the second invited talk to start the afternoon sessions. His presentation entitled *Domain Adaptation of Virtual and Real Worlds for Pedestrian Detection* concerned autonomous driving, and focused on pedestrian (object) detection. Since labelling examples in images is time-consuming and error-prone, he advocated the use of image synthesis to generate examples. Since going from the training data to deployment data requires domain adaptation, he suggested to use active learning to collect a few examples real examples to add to the many virtual examples. The speaker presented many results on real datasets.

2.2 Cost-sensitive learning

The first regular session dealt with cost-sensitive learning.

Ricardo Prudencio presented the first paper, entitled *Cost-sensitive measures of instance hardness*, by Carlos Melo and himself. He considered the cost of misclassifying an instance, averaged over several learning algorithms, with respect to the cost proportion, under three decision threshold choice methods: score-fixed, score-driven, and rate driven.

The second paper of the session was *Classification in Context: Adapting to changes in class and cost distribution* by Peter A. Flach. The talk presented known as well as new results on ROC curves and calibration, in particular with respect to F-measure.

Wouter Duivesteijn presented the third paper *ROCsearch in a Wider Context – A ROC-Guided Search Strategy for Subgroup Discovery and Beyond* by Wouter Duivesteijn, Marvin Meeng and Arno Knobbe. Based on previous work on subgroup discovery published at the 2014 SIAM International Conference on Data Mining, he introduced perspectives for continuous-target subgroup discovery and for exceptional model mining.

2.3 Dataset shift

Meelis Kull presented the first paper, *Patterns of Dataset Shift* by Meelis Kull and Peter Flach. He described a graph notation for modelling different types of dataset shifts, covering known patterns and introducing new patterns as well.

The second paper of the session was *Dataset Shift in a Real-Life Dataset* by Chowdhury Farhan Ahmed, Nicolas Lachiche, Clément Charnay and Agnès Braud. Chowdhury Farhan Ahmed demonstrated that the UCI dataset 'Washington Bike Sharing' exhibits a range of shifts on input as well as output attributes, depending on the seasons, and suggested to use such dataset to compare existing approaches.

Chowdhury Farhan Ahmed also presented the third paper of this session, *Efficient Graph Classification in Shifted Datasets Using Weighted Correlated Feature Selection* by Md. Samiullah, Chowdhury Farhan Ahmed, Anna Fariha and Akiz Uddin Ahmed. The authors proposed a correlation-based graph classification approach

as well as a new diversity capturing measure.

2.4 Multi-instance, multi-dimensional, multi-task and multi-label learning

The third session gathered four papers that generalise the standard classification setting in some way.

The first speaker was Gitte Vanwinckelen, presenting her work with Hendrik Blockeel, *A meta-learning system for multi-instance classification*. They evaluated 14 multiple-instance learning techniques on three dataset domains in order to learn a meta-model. Some interesting conclusions were presented, although they need a larger meta-dataset to generalise.

Adolfo Martínez-Usó presented *Predictive models for multidimensional data when the resolution context changes* by José Hernández-Orallo, Nicolas Lachiche and himself. They considered different aggregation levels in multidimensional data cubes and compared two learning approaches on three datasets.

The third paper of the session was *SymReg-MT: Iterative Multitask Feature Learning Through Weighted Symbolic Regression* by Michael Zwick, Holger Schoner and Ehsan Rezaie. The authors presented an approach to multi-task feature learning that uses weighting functions based on the frequency of occurrences when looking at the models across different tasks.

The last paper of this session, *Multi-label Classification: A Comparative Study on Threshold Selection Methods* by Reem Al-Otaibi, Peter Flach and Meelis Kull was presented by the first author. She evaluated the performance on six datasets of three different methods to set the thresholds: global, label-wise and instance-wise.

2.5 Transfer learning

The session started with the paper *Multi-System Identification for Efficient Knowledge Transfer with Factored Tensor Recurrent Neural Networks* by Sigurd Spieckermann, Siegmund Dull, Steffen Udluft and Thomas Runkler. They proposed a transfer learning approach through parameter transfer to deal with similar dynamical systems and evaluated it on two problems.

Christian Politz presented his paper *Subset based Hilbert space projections for transfer learning*. He proposed two greedy approaches for selecting data samples from the source in order to improve the transfer learning (using a kernel PCA) when target data and source data follow different distributions.

The third paper was *Sequential Decision-Making under Non-stationary Environments via Sequential Change-point Detection* by Emmanuel Hadoux, Aurelie Beynier and Paul Weng. The authors considered reinforcement learning using context detection, defined by a drop of quality. They proposed a new algorithm and evaluated it on three domains.

2.6 Applications

The final session of the day concerned applications.

Niall Twomey presented the first paper of this session, about a work with Peter Flach on *Context Modulation of Sensor Data Applied to Activity Recognition in Smart Homes*. They evaluated whether different subsets of sensors lead to different performances and found that light/door sensors are less valuable than motion sensors on the dataset they considered.

The second paper, *Identifying dominant models when the noise context is known* by Cèsar Ferri, José Hernández-Orallo, Adolfo Martínez-Usó and María José Ramírez-Quintana was presented by the last author. This talk was focussed on an experimental study of several regression algorithms according to the noise level. Assuming that the noise level is known would enable to select the best model.

The last paper was *A L2-Norm Regularized Pseudo-Code for Change Analysis in Satellite Image Time* by Anamaria Radoi and Mihai Datcu. The authors proposed and evaluated a new clustering approach to detect changes in a flow of satellite images.

3. CONCLUSION AND DISCUSSION

The workshop concluded with the general impression that it succeeded as a first workshop on learning models over multiple contexts. First, it was a success mainly because of the number and quality of submissions. Secondly, contributions and participants covered a broad range of related areas. Thirdly, there was a broad consensus that those areas are related and that further efforts to clarify their relationships, similarities and differences would be productive.

The open discussion centred around three issues: 1) domain specificities, 2) support for comparing approaches, and 3) continuation. First, participants noticed that some approaches are domain-specific. For example, domain adaptation often deals with images. It would be worth investigating some cross-over, evaluating approaches of one domain on another application domain. It would need preprocessing data, eg. images, to get a common representation, eg. attribute-value.

Support for sharing data, algorithms and results would help collaborations. Web platforms such as `OpenML.org` or `ClowdFlows.org` were suggested. Another way to promote the comparison of approaches would be to set up a challenge for the next workshop.

Finally, there was a consensus that it would be worthwhile to continue with the workshop. Most people also thought that it should still be associated to a big data mining or machine learning conference: KDD or ECML-PKDD 2015. The possibility of a journal special issue devoted to the topic of model reuse and context change was also considered, possibly after more authors have had a chance to present their work at the second workshop.

Acknowledgements