Final Program
IAPR Workshop on Partially Supervised Learning - PSL 2011
Sept 15-16, 2011, University of Ulm

SEPT 15, 2011

9:00 *REGISTRATION*
9:30 *OPENING*

SESSION I - Ensembles and combined classifiers
Chair: tba

9:40 Yann Soullard and T. Artieres (University Pierre and Marie Curie, Paris, France)
Iterative Refinement of HMM and HCRF for Sequence Classification

We propose a strategy for semi-supervised learning of Hidden-state Conditional Random Fields (HCRF) for signal classification. It builds on simple procedures for semi-supervised learning of HMMs and on strategies for learning a HCRF from a trained HMM system. The algorithm learns in parallel a generative system based on Hidden Markov models and a discriminative one based on HCRFs which are iteratively refined.

10:05 Ilaria Castelli and Edmondo Trentin (University of Siena, Italy)
Supervised and Unsupervised Co-Training of Adaptive Activation Functions in Neural Nets

In spite of the nice theoretical properties of mixtures of logistic activation functions, standard feedforward neural network with limited resources and gradient-descent optimization of the connection weights may practically fail in several, difficult learning tasks. Such tasks would be better faced by relying on a more appropriate, problem-specific basis of activation functions. The paper introduces a connectionist model which features adaptive activation functions. Each hidden unit in the network is associated with a specific pair ($\mathbf{f}(\cdot)$, $\mathbf{p}(\cdot)$), where $\mathbf{f}(\cdot)$ (the very activation) is modeled via a specialized neural network, and $\mathbf{p}(\cdot)$ is a probabilistic measure of the likelihood of the unit itself being relevant to the computation of the output over the current input. While $\mathbf{f}(\cdot)$ is optimized in a supervised manner (through a novel backpropagation scheme of the target outputs which do not suffer from the traditional phenomenon of "vanishing gradient" that occurs in standard backpropagation), $\mathbf{p}(\cdot)$ is realized via a statistical parametric model learned through unsupervised estimation. The overall machine is implicitly a co-trained coupled model, where the topology chosen for learning each $\mathbf{f}(\cdot)$ may vary on a unit-by-unit basis, resulting in a highly non-standard neural architecture.

10:30 *COFFEE BREAK*

INVITED SESSION
Chair: tba

11:00 **Zhi-Hua Zhou (Nanjing University, China)**
Unlabeled Data and Multiple Views

12:15 *LUNCH*

INVITED SESSION
Chair: tba

14:15 **Catrin O. Plumpton (Bangor University, UK)**
Online Semi-Supervised Ensemble Updates for fMRI Data

15:15 *COFFEE BREAK*

POSTERSESSION

J. Esparza, S. Scherer, F. Schwenker (University of Ulm, Germany)
Studying Self- and Active-Training Methods for Multi-Feature Set Emotion Recognition

C. I. Cooper, R. Kilmer (Fort Lewis College, Walden University, USA)
Using Self Organizing Maps to Find Good Comparison Universities

E. Trentin, L. Lusnig, F. Cavalli (University of Siena, Italy)
Comparison of Combined Probabilistic Connectionist Models in a Forensic Application

M. Kächele, M. Schels, D. Hrabal, S. Walter, H. Traue, F. Schwenker (University of Ulm)
Classification of emotional states in a WOZ Scenario exploiting labeled and unlabeld bio-physiological data

D. A . Popescu (University of Pitesti, Romania)
Sink web pages in Web Application

I. Castelli, E. Trentin (University of Siena, Italy)
Semi-Unsupervised Weighted Maximum-Likelihood Estimation of Joint Densities for the Co-Training of Adaptive Activation Functions

19:00 *WORKSHOP DINNER at Gasthaus „Drei Kannen“*

Final Program
IAPR Workshop on Partially Supervised Learning - PSL 2011
Sept 15-16, 2011, University of Ulm

SEPT. 16, 2011

SESSION II – Algorithms
Chair: tba

9:15 Stefan Fausser, Friedhelm Schwenker (University of Ulm, Germany)
Semi-Supervised Kernel Clustering with Sample-to-Cluster Weights

Collecting unlabelled data is often effortless while labelling them can be diffcult. Either the amount of data is too large or samples cannot be assigned a specific class label with certainty. In semi-supervised clustering the aim is to set the cluster centres close to their label-matching samples and unlabelled samples. Kernel based clustering methods are known to improve the cluster results by clustering in feature space. In this paper we propose a semi-supervised kernel based clustering algorithm that minimizes convergently an error function with sample-to-cluster weights. These sample-to-cluster weights are set dependent on the class label, i.e. matching, not-matching or unlabelled. The algorithm is able to use many kernel based clustering methods although we suggest Kernel Fuzzy C-Means, Relational Neural Gas and Kernel K-Means. We evaluate empirically the performance of this algorithm on two real-life dataset, namely Steel Plates Faults and MiniBooNE.

9:40 Kazuki Yoshiyama, Akito Sakurai (Keio University, Japan)
Manifold-Regularized Minimax Probability Machine

In this paper we propose Manifold-Regularized Minimax Probability Machine, called MRMPM. We show that Minimax Probability Machine can properly be extended to semi-supervised version in the manifold regularization framework and that its kernelized version is obtained for non-linear case. Our experiments show that the proposed methods achieve results competitive to existing learning methods, such as Laplacian Support Vector Machine and Laplacian Regularized Least Square for publicly available datasets from UCI machine learning repository.

10:05 Marco Loog (Delft University of Technology, The Netherlands)
Semi-Supervised Linear Discriminant Analysis Using Moment Constraints

A semi-supervised version of a classical classifier, i.e., linear discriminant analysis, is presented. As opposed to the lion's share of current approaches to semi-supervision, no assumptions on the data distribution are made, apart from the ones explicitly or implicitly present in standard supervised linear discriminant analysis. Our approach exploits the fact that the parameters that are to be estimated in linear discriminant analysis fulfill particular relations that link label-dependent with label-independent quantities. In this way, the latter type of parameters, which can be estimated based on unlabeled data, impose constraints on the former and lead to a reduction in variability of the label dependent estimates. As a result, the performance of our semi-supervised linear discriminant is expected to improve over that of its regular supervised counterpart and typically does not deteriorate with increasing numbers of unlabeled data.

10:30 *COFFEE BREAK*

INVITED SESSION
Chair: tba

11:00 **Stefan Scherer (Trinity College Dublin, Ireland)**
How Partially Supervised Learning can facilitate and enhance user state analysis in naturalistic HCI

12:15 *LUNCH*

SESSION III – Applications
Chair: tba

14:15 Matthias Hillebrand, Christian Wöhler, Ulrich Kreßel, Franz Kummert
(Daimler AG Ulm, University of Dortmund, University of Bielefeld, Germany)
Semi-supervised Training Set Adaption to Unknown Countries for Traffic Sign Classifiers

Traffic signs in Western European countries share many similarities but also can vary in colour, size, and depicted symbols. Statistical pattern classification methods are used for the automatic recognition of traffic signs in state-of-the-art driver assistance systems. Training a classifier separately for each country requires a huge amount of training data labelled by human annotators. In order to reduce these efforts, a self-learning approach extends the recognition capability of an initial German classifier to other European countries. After the most informative samples have been selected by the confidence band method from a given pool of unlabelled trafic signs, the classifier assigns labels to them. Furthermore, the performance of the self-learning classifier is improved by incorporating synthetically generated samples into the self-learning process. The achieved classification rates are comparable to those of classifiers trained with fully labelled samples.

14:40 Andreas Müller, Sven Behnke (University of Bonn, Germany)
Multi-Instance Methods for Partially Supervised Image Segmentation

In this work, we explore the application of multi-instance learning algorithms to the task of partially supervised image segmentation. Multi-Instance learning is a natural formulation for image classification and has been successfully applied in this task. We propose to go a step further and apply multi-instance learning to the task of object class segmentation in natural images. In object class segmentation, the goal is to create a pixel-wise labeling of an input image into one of several semantic classes. Multi-class image segmentation receives much attention in the computer vision community at the moment. To our knowledge, all previous methods in the field use strong supervision, meaning manual pixel-wise annotation of training images. This approach does not scale to larger datasets, especially if one expects consistency and quality in the segmentations. In this work, we investigate the use of multi-instance learning to obtain multiclass image segmentations using ground truth labeling only on the image level. We focus on the multi-class, single label setup, where each image is assigned one of multiple classes. We formulate multi-class image segmentation as a multi-instance learning problem by considering each image as a bag of overlapping candidate segments.

15:05 *COFFEE BREAK*

SESSION IV – Algorithms
Chair: tba

15:35 Simon Smith Bize, J. Michael Herrman (University of Edinburgh, UK)
Homeokinetic Reinforcement Learning

Reinforcement learning includes establishing of a relation between a value function and a policy. In autonomous robots, this relation is often unavailable but can be revealed locally by modulating the motor command by probing actions. For robots with many degrees of freedom, this type of exploration becomes inefficient such that is is an interesting option to leave the choice of the probing actions to a self-organising auxiliary controller. In contrast to the approach of guided self-organisation where a self-organising process is merely modulated, we suggest here the more direct approach to optimise the exploration in reinforcement learning. The approach is illustrated by a simple standard learning problem and applied to the generation of walking behaviour in an hexapod robot.

16:00 Ludwig Lauser, Florian Schmid, Hans A. Kestler (University of Ulm, Germany)
On the utility of partially labeled data for classification of microarray data

Microarrays are standard tools for measuring thousands of gene expression levels simultaneously. They are used in the classification process of tumor tissues (i.e., "benign" vs "malign"). In this setting a collected set of samples often consists only of a few dozen data points. Common approaches for classifying such data are supervised. They exclusively use categorized data for training a classification model. Restricted to a relatively small number of samples, these algorithms are affected by overfitting and often lack a good generalization performance. An implicit assumption of supervised methods is that only labeled training samples exist. This assumption does not always hold in real life studies. Here often additional unlabeled samples are available. In some settings collected data points can be unlabeled for several years (i.e., "early relapse" vs."late relapse"). Other classification approaches, such as semi-supervised or transductive algorithms, are able to handle partially labeled data. These methods incorporate information from unlabeled samples during their training phase. Extended by these additional samples the available amount of training data is larger for these learning schemes. This might also reduce overfitting effects. Candidates of this category are the self-learning algorithm of Yarowsky or the mincut strategy of Blum and Chawla. The focus of our study is the usage of semi-supervised and transductive algorithms as "early prediction tools" for datasets of low cardinality. We empirically investigate the generalization performance and robustness of semi-supervised or transductive algorithms in settings with varying ratio of labeled to unlabeled examples.

16:25 *WORKSHOP CLOSING*