Learning with Partial Supervision

Abdelhamid Bouchachia
University of Klagenfurt, Austria

INTRODUCTION

Recently the field of machine learning, pattern recognition, and data mining has witnessed a new research stream that is *learning with partial supervision* -LPS- (known also as *semi-supervised learning)*. This learning scheme is motivated by the fact that the process of acquiring the labeling information of data could be quite costly and sometimes prone to mislabeling. The general spectrum of learning from data is envisioned in Figure 1. As shown, in many situations, the data is neither perfectly nor completely labeled.

LPS aims at using available labeled samples in order to guide the process of building classification and clustering machineries and help boost their accuracy. Basically, LPS is a combination of two learning paradigms: supervised and unsupervised where the former deals exclusively with labeled data and the latter is concerned with unlabeled data. Hence, the following questions:

- Can we improve supervised learning with unlabeled data?
- Can we guide unsupervised learning by incorporating few labeled samples?

Typical LPS applications are medical diagnosis (Bouchachia & Pedrycz, 2006a), facial expression recognition (Cohen et al., 2004), text classification (Nigam et al., 2000), protein classification (Weston et al., 2003), and several natural language processing applications such as word sense disambiguation (Niu et al., 2005), and text chunking (Ando & Zhangz, 2005).

Because LPS is still a young but active research field, it lacks a survey outlining the existing approaches and research trends. In this chapter, we will take a step towards an overview. We will discuss (i) the background of LPS, (iii) the main focus of our LPS research and explain the underlying assumptions behind LPS, and (iv) future directions and challenges of LPS research.

BACKGROUND

LPS is about devising algorithms that combine labeled and unlabeled data in a symbiotic way in order to boost classification accuracy. The scenario is portrayed in Fig.2 showing that the combination can mainly be done in two ways: active/passive pre-labeling, or via 'pure' LPS (Fig. 4). We try to draw a clear picture about these schemes by means of an up-to-date taxonomy of methods.

Active and Passive Pre-Labeling

Pre-labeling aims at assigning a label to unlabeled samples (called queries). These samples are used together with the originally labeled samples to train a fully supervised classifier (Fig. 3). "Passive" pre-label-
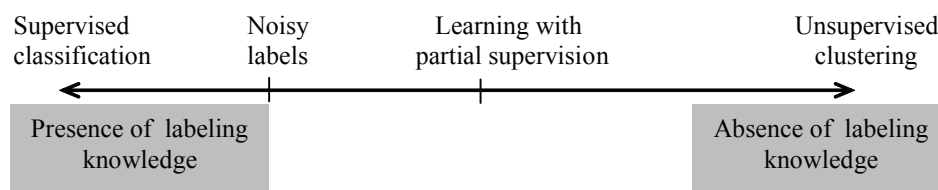
Figure 1. Learning from data spectrum

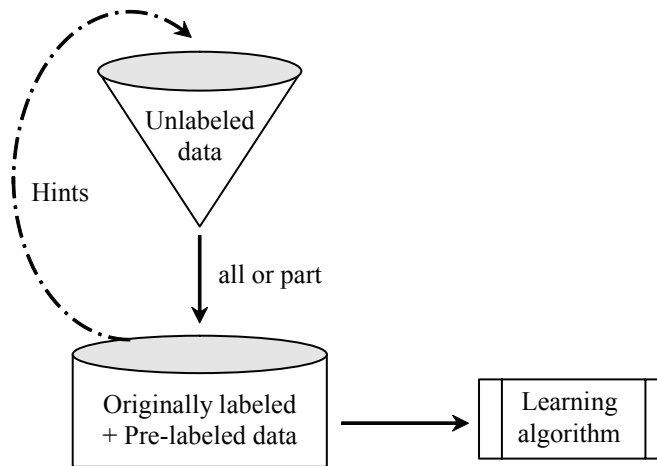*Figure 2. Combining labeled and unlabeled data*
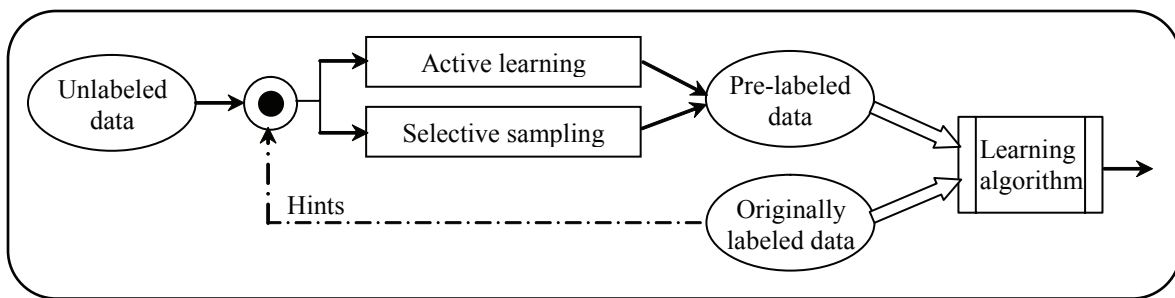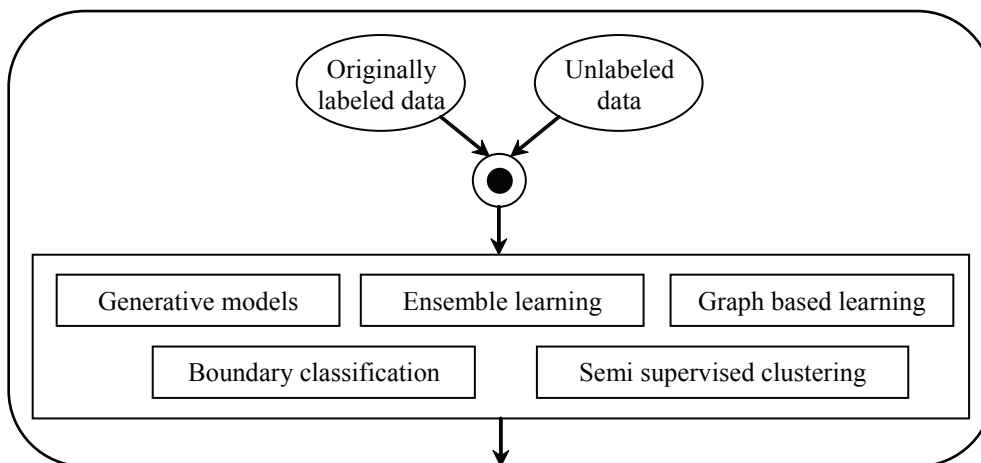

Figure 3. Pre-labeling approaches

Figure 4. Combining labeled and unlabeled data

Related Content

Distance-Based Methods for Association Rule Mining

Vladimír Bartík (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 689-694).*

www.irma-international.org/chapter/distance-based-methods-association-rule/10895

Data Mining for Fraud Detection System

Roberto Marmo (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 411-416).*

www.irma-international.org/chapter/data-mining-fraud-detection-system/10853

Wrapper Feature Selection

Kyriacos Chrysostomou (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 2103-2108).*

www.irma-international.org/chapter/wrapper-feature-selection/11110

The Issue of Missing Values in Data Mining

Malcolm J. Beynon (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 1102-1109).*

www.irma-international.org/chapter/issue-missing-values-data-mining/10959

Association Rules and Statistics

Martine Cadot, Jean-Baptiste Majand Tarek Ziadé (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 94-97).*

www.irma-international.org/chapter/association-rules-statistics/10804