

Multilabel classification as an important research field in machine learning arises naturally from many real world applications. For example, in document classification, a research article can be categorized into “science”, “drug” and “genomics” at the same time. The goal of multilabel classification is to make reliable prediction on multiple output labels for a given input example. As multiple interdependent output labels can be “on” and “off” simultaneously, the central problem in multilabel classification is to exploit the label correlation in order to make accurate prediction. Compared to the previous flat multilabel classification approach which treats the multiple labels as a flat vector, structured output learning relies on an output graph connecting multiple labels in order to model the label correlation in a comprehensive manner. The main question studied in the thesis is how to tackle multilabel classification through structured output learning. Within this scope we discuss several subproblems.

The thesis starts with extensive review on the topic of classification learning covering both single-label and multilabel classification settings. The first problem that we address is how to solve the multilabel classification problem when output graph is observed as a-priori. We revisit several well established structured output learning algorithms and study the network response prediction problem within the context of social network analysis. We realize that the current structured output learning algorithms rely on the output graph to gain representation power of label dependency. Therefore, the second problem that we address is how to use structured output learning when there is no pre-established output graph. More specifically, we examine the potential of learning on a set of random output graphs when the “real” one is hidden. This problem is relevant as in most multilabel classification problems there does not exist any output graph that reveals the label dependency. The third problem that we address is how to analyze the proposed learning algorithms in a theoretical manner. Especially, we want to explain the behavior of the proposed models and to study the generalization error.

The main contributions of the thesis are the new learning algorithms that widen the applicability of structured output learning. For the problem with observed structure, the proposed algorithm “SPIN” is able to predict a directed acyclic graph from an observed underlying network. For general multilabel classification problems without pre-established output graph, we proposed several learning algorithms that aim to combine many structured output learners built on random output graphs. In addition, we develop a joint learning and inference framework which is based on the Max-Margin learning over a random sample of spanning trees. The theoretic analysis also guarantees the generalization error of the proposed methods.