Multi-Label Classification

Jesse Read  
*Universidad Carlos III, Spain*

Albert Bifet  
*Yahoo! Research Barcelona, Spain*

**INTRODUCTION**

Humans naturally associate an object with more than one concept, i.e., *label* (also known as a category, tag, or genre). Adding labels to data collections can greatly facilitate retrieval and organization. This is a vital task with the ever-increasing collections of data being created and accessed each day. In 2010 it was estimated that we were creating the same quantity of data every two days, as we created from the dawn of time up until 2003. Multi-label classification is a way to learn label multi-label associations and, importantly, assign labels to new data automatically.

**BACKGROUND**

Multi-label classifications come naturally. For example, a movie can be assigned to both genres “Action” and “Comedy.” In the physical world, this often involves physical duplication, for example, a video store has to put copies of such a movie in two different sections -- “Action” and “Comedy” -- or create a single-category (“Action-Comedy”). In a digital environment there are no such restrictions, and it is likely for this reason that there has been a rapid rise in the popularity of the multi-label concept in everyday applications. Consider how the once-familiar ‘folder’ metaphor is being replaced by the label/tag term in many everyday applications: email, picture, document, and media collections, and so forth.

The typical goal of multi-label learning is to learn a model to predict or recommend labels for data points automatically, and thus reduce or even eliminate the need for time-consuming manual labeling of data and document collections. Such collections may include e-mails and text documents, images, audio and video collections, or even certain biological applications such as where genes can be associated with multiple functions. For a detailed introduction and review of multi-label classification, see for example (Tsoumakas, G., Katakis, I., & Vlahavas, 2010).

Multi-label classification has borrowed heavily from the already-existing domain of traditional single-label classification, of assigning a single ‘class’ to each data element. Methods can be grouped roughly into two categories:

**Problem/Data Transformation**

In this approach, the multi-labeled data is transformed into one or a series of single-label problems. Standard off-the-shelf single-label classifiers can then be applied. A typical approach is to create one binary problem for each label (to predict if the label is relevant or not), as in (Read et al., 2011), or a multi-class problem where combinations of labels are represented as atomic mutually-exclusive classes (“Action-Comedy,” in the movie example, would be considered a