Chapter 4
Multi–Label Classification Method for Multimedia Tagging

Aiyesha Ma
Oakland University, USA

Ishwar Sethi
Oakland University, USA

Nilesh Patel
Oakland University, USA

ABSTRACT

Community tagging offers valuable information for media search and retrieval, but new media items are at a disadvantage. Automated tagging may populate media items with few tags, thus enabling their inclusion into search results. In this paper, a multi-label decision tree is proposed and applied to the problem of automated tagging of media data. In addition to binary labels, the proposed Iterative Split Multi-label Decision Tree (IS-MLT) is easily extended to the problem of weighted labels (such as those depicted by tag clouds). Several datasets of differing media types show the effectiveness of the proposed method relative to other multi-label and single label classifier methods and demonstrate its scalability relative to single label approaches.

INTRODUCTION

The retrieval of multimedia information for search queries can be based on two basic approaches: similarity of multimedia artifacts, and similarity of associated keywords or tags. Although the bulk of content-based retrieval work falls within the first category, providing an example document for the query may be difficult or impossible. Thus

the second approach allows the search to proceed without an example document, but relies upon the association of keywords or tags to the multimedia item. Many websites allow the user community to supply these tags, thus enabling easier retrieval. For example, a YouTube user may add tags when adding a video, while the entire community may manually specify content tags on Flickr. Not all users or communities are equally ambitious with their tagging, however. Thus, automatic assignment or suggestions of tags for new or untagged

multimedia documents may help to improve the retrieval of multimedia content.

Existing tag data can be used for training classifiers to enable the automatic assignment of keywords or tags (Snoek & Worring, 2009). A multimedia document is likely to have many tags associated with it, but traditional classifiers make assignments to disjoint categories. Although a classifier could be created for each potential label, this approach does not scale to large numbers of tags. Multi-label classifiers, however, assign a set of labels using a single classifier. These classifiers exploit data similarities between tags, such that a single multi-label classifier may perform better than a set of single label classifiers, (Ueda & Saito, 2003; Rousu et al., 2004; Blockeel et al., 2006; Vens et al., 2008), in addition to being more scalable.

To this end, we developed a multi-label decision tree and applied it to the problem of automated tagging. The basic idea consists of partitioning the set of labels into two groups at each node. Using the two groups a node level decision is created using a binary decision method such as SVM or Information Gain. After constructing the tree, the leaf node frequencies are considered scores and labels are assigned by thresholding. The proposed method was applied to the problem of tagging

emotions in music sound clips; this paper slightly expands the results in our earlier work, (Ma et al., 2009), by varying the training-validation ratios as well as comparing the multi-label classifier methods to single-label classifiers. This paper further demonstrates the proposed multi-label decision tree by applying it to the problem of tagging photographs with scene descriptors.

Some websites with community tagging depict tag clouds (see Figure 1). Thus, instead of the binary input of inclusion and exclusion of a tag, each tag has a weight associated with it that generally indicates the proportion of users who included that particular label when tagging the content. Although these weighted tags could be converted to the binary problem of inclusion and exclusion, these weights may provide additional information to the classifier. To the best of our knowledge no multi-label classifier methods have addressed this weighted labeling problem.

Since the premise of our multi-label decision tree is to partition the set of labels into two groups, it is easily extended to this weighted multi-label problem: rather than using a binary distance measure in the partition clustering, a real valued distance measure is used, and the performance measure used to determine the amount of pruning is similarly changed from a binary evaluation

Figure 1. Example of a Tag Cloud: The Beach Boys, from Last.FM
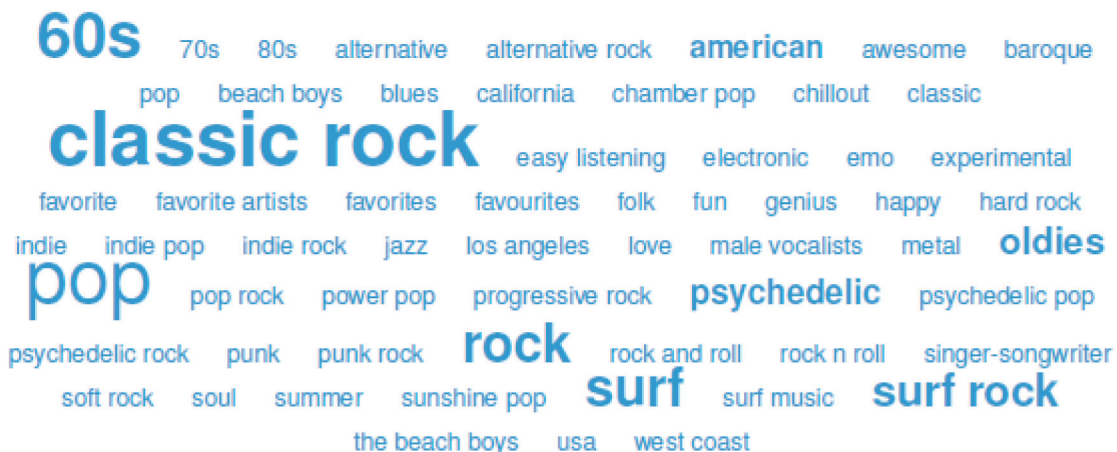
Related Content

Location-Aware Caching for Semantic-Based Image Queries in Mobile AD HOC Networks

Bo Yangand Manohar Mareboyana (2012). *International Journal of Multimedia Data Engineering and Management (pp. 17-35).*

www.irma-international.org/article/location-aware-caching-semantic-based/64629

Screenspace

K. H. Jones (2008). *Multimedia Technologies: Concepts, Methodologies, Tools, and Applications (pp. 721-730).*

www.irma-international.org/chapter/screenspace/27117

Temporal Models and Their Applications in Multimedia Information Retrieval

Balakrishnan Prabhakaran (2001). *Design and Management of Multimedia Information Systems: Opportunities and Challenges (pp. 345-379).*

www.irma-international.org/chapter/temporal-models-their-applications-multimedia/8112

Exploring Different Optimization Techniques for an External Multimedia Meta-Search Engine

Kai Schlegel, Florian Stegmaier, Sebastian Bayerl, Harald Koschand Mario Döller (2012). *International Journal of Multimedia Data Engineering and Management (pp. 31-51).*

www.irma-international.org/article/exploring-different-optimization-techniques-external/75455

Ontology Instance Matching based MPEG-7 Resource Integration

Hanif Seddiquiand Masaki Aono (2010). *International Journal of Multimedia Data Engineering and Management (pp. 18-33).*

www.irma-international.org/article/ontology-instance-matching-based-mpeg/43746