UNED at MediaEval 2011: Can Delicious help us to improve automatic video tagging?

J. Cigarran, V. Fresno, A. García-Serrano, D. Hernández-Aranda, R. Granados,
{juanci, vfresno, agarcia, daherar, rgranados}@lsi.uned.es
NLP & IR Group, UNED, Madrid

ABSTRACT

In this paper we present the second participation of the NLP&IR group at UNED in the MediaEval Genre Tagging Task. This categorization task was carried out applying an Information Retrieval (IR) approach considering the video collection's textual data and query expansion techniques. The results show that the combination of social tags and language models is useful to perform query expansion.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*Search Process*

Keywords

tag prediction, video indexing, multimedia information retrieval, query expansion

1. INTRODUCTION

In this paper we introduce a new approach for genre tagging task within the MediaEval initiative [3]. We are only focused on textual information and our contribution is mainly based on exploiting information extracted from social tagging sites like Delicious. Until now, most of the automated classifiers rely on the content of the videos, both textual, audio and video information. Nonetheless, the lack of representative data within many videos makes this classification complex. In some cases, it may not be feasible to obtain an accurate representation for video resources in a genre tagging task. Transcripts are not usually representative with respect to specific tags as, e.g., "comedy" or "tutorial". As a mean to solve these issues, social systems can provide an easier and cheaper way to obtain useful information.

First, we will introduce the related work, then we will present the experimental setup and we will focus on the query expansion techniques applied. Then, we will describe the submited runs and, finally, we will discuss the results.

2. RELATED WORK

Social annotations have been widely used for the sake of information management tasks. Before [8], there was little work dealing with the analysis of the applicability and usefulness of social tags for resource classification and information retrieval tasks. [1] presented a study of the characteristics of social annotations provided by end users, in order to determine their usefulness for web page classification. In [6] they studied the usefulness of social tags as a complementary source for improving the classification of academic conferences into corresponding topics. With regard to the classification of resources other than web pages, in [4] they present a comparison of tags annotated on books and their Library of Congress subject headings and in [7] the authors present a method to classify products from Amazon into their corresponding categories using social tags.

3. SYSTEM OVERVIEW

Our tagging system applies an IR approach which uses the candidate labels as queries. The IR module applies a BM25F ranking function [5] to calculate the similarities. This function has shown a god performance on those scenarios where data is represented and indexed on different fields. In our particular case, we have indexed the MediaEval collection considering the different textual data (i.e. trancripts, metadata and tags) as separate fields. As the use of standalone labels to query the system could drastically reduce the recall performance we decided to expand them using three different approaches: a) **Query expansion using Delicious**. This approach uses Delicious tags to expand each label. We used the DeliciousT140 Dataset [1] which is made up by 144.574 unique URLs, all of them with their corresponding social tags retrieved from Delicious on June 2008. This set of documents is annotated with 67.104 different social tags and it has also useful frequency information. Firstly, we processed the dataset to find those tags co-occurring with each targeted label. For each label we generated a set of Delicious tags ordered by use frequency. The size of these sets was between 400 and 2000 co-occurring tags and they were characterized by a long tail where only a few of them had consensus between the most Delicious users (i.e. highest frequencies). Thus, the query expansion process needed to filter and weight the obtained set in order to reduce the final recall. For each Delicious tag set we selected only those with a frequency over the 10% of the highest tag frequency. Finally, and to provide more value to higher frequency tags and also to the targeted label, we weighted the expanded query terms. b) **Query expansion using KLD**. As the labels set was the same for the development and also for the test phases, this approach expands each targeted label with

[1] http://nlp.uned.es/social-tagging/delicioust140/

MediaEval 2011 Workshop, September 1-2, 2011, Pisa, Italy

specific vocabulary extracted from the development set. We used the Kullback-Leibler divergence (KLD) [2] to rank the terms occurring on the relevant video items for each label with respect to the whole video collection. To do this process we used the transcripts, the metadata and the user tags and we considered the subset of relevant videos as a large document to compute the divergence. As in the previous approach, a filtering and weighting process was needed, we selected only those terms with a weight over the 33% of the highest term weight due to there were no significant differences (i.e., there was no long tail). Finally, we also weighted the query terms in order to provide more value to the targeted label and higher weighted terms. c) **Combination of Delicious and KLD**. The idea behind is to filter those Delicious tags that, although they could be adequate for a big collection such as Delicious, they are out of the scope of the MediaEval collection. Thus, we used KLD to drive the query term expansion. More specifically, we intersected both sets of terms and Delicious tags obtained from the previous approaches and then we applied a frequency based selection and query terms weighting to build the query.

4. SUBMITTED RUNS

We submitted five different runs combining the different approaches presented in section 3:

Run 1. We queried the retrieval system using a subset of Delicious tags. The video collection was represented considering only the video transcripts.

Run 2. We also used Delicious as the query expansion method but in this case, the video collection was represented and indexed using the metadata and also the video transcripts.

Run 3. We evaluated the same query expansion approach but considering only the video metadata and the provided social tags in the representation and indexing phase.

Run 4. We queried the retrieval system using the KLD expansion approach. The video collection was represented using the video metadata and the provided social tags.

Run 5. We combined both query expansion approaches, KLD and Delicious tags . Again, in this approach the video collection was represented using the video metadata and the social tags.

5. ANALYSIS OF RESULTS

Table 1 shows the results obtained by our runs according to Mean Average Precission (MAP), which was the official measure for this task. Runs 1, 2 and 3 show how the representation of the video collection using combinations of different data impacts in the retrieval process. In these cases, the representation of the video collection using its transcripts (runs 1 and 2) get the worst results and it is not suitable for retrieval using Delicious query expansion. On the other hand, runs 3 and 4 show a better performance, which indicates that the representation of the video collection using semantic and social data such as video metadata and tags is more adequate to the proposed query expansion approaches. Finally, run 5 shows that the combination of Delicious social tags and specific vocabulary extracted using KLD performs better than any of the previous runs. This combination helps to adjusts the candidate set of query terms to the language model of each targeted label and improves the final retrieval performance.

RUN	MAP
1	0.1103
2	0.1111
3	0.1850
4	0.1836
5	0.2070

Table 1: Results of the submitted runs

6. CONCLUSIONS

In this paper we have presented an IR based approach for automatic genre tagging. We have considered only the textual data associated with the video collection and we have described several techniques to perform query expansion based on Delicious social tags and language models. According to the results, the use of social tags and also its filtering using specific vocabulary from the collection improves the final retrieval results.

Acknowledgments

This work has been partially supported by the BUSCAMEDIA Project (CEN-20091026).

7. REFERENCES

[1] D. Godoy and A. Amandi. Exploiting the social capital of folksonomies for web page classification. In *Software Services for E-World*, volume 341 of *IFIP Advances in Information and Communication Technology*, pages 151–160. Springer, 2010.

[2] S. Kullback and R. A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22(1), 1951.

[3] M. Larson, M. Eskevich, R. Ordelman, C. Kofler, S. Schmiedeke, and G. Jones. Overview of MediaEval 2011 Rich Speech Retrieval Task and Genre Tagging Task. In *MediaEval 2011 Workshop*, Pisa, Italy, September 1-2 2011.

[4] C. Lu, J.-r. Park, and X. Hu. User tags versus expert-assigned subject terms: A comparison of librarything tags and library of congress subject headings. *Journal of Information Science*, 36(6):763–779, 2010.

[5] J. Pérez-Iglesias, J. R. Pérez-Agüera, V. Fresno, and Y. Z. Feinstein. Integrating the Probabilistic Models BM25/BM25F into Lucene. *CoRR*, abs/0911.5046, 2009.

[6] J. Xia, K. Wen, R. Li, and X. Gu. Optimizing academic conference classification using social tags. *Computational Science and Engineering, IEEE International Conference on*, 0:289–294, 2010.

[7] Z. Yin, R. Li, Q. Mei, and J. Han. Exploring social tagging graph for web object classification. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '09, pages 957–966, New York, NY, USA, 2009. ACM.

[8] A. Zubiaga. *Harnessing Folksonomies for Resource Classification*. PhD thesis, Madrid, Spain, 2011. Adviser-Victor Fresno and Adviser-Raquel Martinez.