Comparative Evaluations of a Hierarchical Categorization of Search Results based on a Granular View of Domain Ontologies

Silvia Calegari, Fabio Farina[*] and Gabriella Pasi
DISCo, Università di Milano–Bicocca, Milano, Italy

Keywords: Granular Domain Ontology, Search Results Categorization, Evaluations.

Abstract: The aim of this paper is to evaluate the effectiveness of a categorization approach of search results based on the use of domain ontologies with respect to the application of standard single-label and multi-label classification algorithms. In particular, the approach in (Calegari et al., 2011) is considered, where the categorization process is performed thanks to the adoption of a taxonomy of information represented as a granular view of a domain ontology.

1 INTRODUCTION

The huge amount of information on the Web is always in continuous overwhelming. A drawback of this situation is that for people is more difficult to discover out information that could satisfy their information needs as by the evaluation of a query produced by a search engine several pages of results can be obtained. Indeed, people often use search engines to formulate queries related to their specific interests (such as professional activities or hobbies); and they spend a lot of time to focusing the attention on relevant search results. The identification of relevant search results is becoming a time consuming activity; to support the user during his/her searches on the Web, in the literature several solutions have been proposed such as personalization issues (Daoud et al., 2008; Ma et al., 2007; Shahabi and Chen, 2003; Calegari and Pasi, 2010; Baeza-Yates and Maarek, 2011; Teevan et al., 2005; Calegari and Pasi, 2008), relevance feedback approaches (Salton and Buckley, 1990; Lv and Zhai, 2010; Halpin and Lavrenko, 2011; Ruthven and Lalmas, 2003), etc.

The research reported in this paper is related to the problem of categorizing the results produced by search engines in response to users' queries. The task is to help users in easily identifying the search results that could satisfy at best their information needs. Twofold is the advantage of using a categorization method for the Web: (1) a set of search results is grouped into one or more categories, and (2) the label of each category is set with a suited semantic that expresses the meaning of the categorized search results, respectively. This way, a user can identify the relevant search results faster.

In the literature, both unsupervised and supervised techniques of categorization have been proposed (Carpineto et al., 2009; Sebastiani, 2002). The focus of this paper is on unsupervised approaches that make use of an external reference knowledge, generally a taxonomy, to associate each search result with one or more categories of the considered resource. For example, in (Ren et al., 2009) a general purpose ontology to couple each search result with one category of the ontology (single-label categorization) is used. Instead, in (Calegari et al., 2011) a hierarchical multi-label categorization approach based on the use of granular views of a domain ontology is proposed. The usage of a domain ontology to categorize Web search results allows to support users in easily identifying their relevant search results than a general purpose ontology. As (1) it avoids the problem of categorizing search results in ambiguous categories, (2) it limits the number of categories used during the categorization process by reducing the time of selection of the right category, and (3) people often use search engines to formulate queries on specific interests that refer to topical domains.

However, to define effective categorization algorithms based on a domain ontology is not an easy task as ontologies are complex structures. In (Calegari and Ciucci, 2010) a method to generate a granular view of an ontology is proposed. A granular view is a compact representation of an ontology where the

[*]Now at Consortium GARR, the Italian NREN.

Calegari S., Farina F. and Pasi G..
Comparative Evaluations of a Hierarchical Categorization of Search Results based on a Granular View of Domain Ontologies.
DOI: 10.5220/0004625003610366
In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development* (KEOD-2013), pages 361-366
ISBN: 978-989-8565-81-5

ontology concepts are grouped into coarser granules based on shared properties of the original concepts. Then, a granular view allows to manage the concepts linked with different properties in order to define a new subsumption relation among the analyzed concepts according to common properties that they share. The obtained granular ontology based on the new subsumption relation is named *granular view*, and it has less nodes (concepts) than the original domain ontology; as a consequence a categorization process based on this simplified representation is faster and simpler.

The objective of this paper is to evaluate the categorization method proposed in (Calegari et al., 2011), that can be applied with different levels of accuracy based on which information granules are selected in the granular view of the reference domain ontology to assess its efficacy. In particular, two kinds of categorization may be performed, i.e. single-label and multi-label. To this goal, we extend the approach defined in (Calegari et al., 2011) to also guarantee a single-label categorization. To assess the effectiveness of the single-label categorization, in this paper, some comparative evaluations have been performed with respect to standard single-label classifiers i.e., the Multinomial Naïve Bayes, the linear SVM (i.e. LibLinear), and the quadratic SVM (i.e. LibSVM, classifiers). Moreover, the results of another kind of evaluation are reported in this paper: the effectiveness of the multi-label categorization has been evaluated by comparing the search results produced by using the original domain ontology with those produced by using its granular view.

The paper is organized as follows: Section 2 gives the methodology used for the multi-label and the single-label categorization based on a granular view, Section 3 and Section 4 present the details concerning the metrics used for the evaluations, and the data of the experiments, respectively. Finally, in Section 5 some conclusions and future activities are stated.

2 CATEGORIZATION OF SEARCH RESULTS BASED ON A GRANULAR VIEW

In this section, we shortly sketch the categorization method (Calegari et al., 2011) used to associate each search result with one or more topical granules belonging to the granular view, named multi-label categorization. Here, we extend the considered categorization method with the case of single-label categorization. The new property allows to the system to associate each search result only with a topical granule of the granular view. This way, the categorization method can be compared with standard single-label classifiers to test its effectiveness with approaches recognized as milestone evaluators in the literature (see Section 3 and Section 4).

The notion of a granular view of a domain ontology **O** is formally defined as a pair:

$$\mathbf{G_O} = \{\mathcal{G}, \mathbf{R_{IS-A}}\},$$

where $\mathcal{G}$ is a set of granules, and $\mathbf{R_{IS-A}}$ is the IS-A subsumption relation defined on the set of granules $\mathcal{G}$. To generate a granular view of a domain ontology means to group into granules the entities of **O** that share some properties and properties value. From **O** only the instances linked by the IS-A relation and the properties defined on them are considered. More formal details concerning how the granular view is obtained are reported in (Calegari and Ciucci, 2010; Calegari et al., 2011) as this aspect is out of topic with respect to the aim of this paper.

Multi-label Categorization. Let *Res* the set of search results, then the association of a search result $R_i \in Res$ with one or more granules $g_k \in \mathcal{G}$ is performed by two subsequent steps:

1. "Search results conceptual indexing". Both title and snippet of R_i are indexed via the controlled vocabulary constituted by $\mathcal{G}$; we denote by $Rep(R_i)$ the set of the representative granules extracted from title and snippet
2. "Association of search results R_i with granule g_k". For each granule g_k in $Rep(R_i)$ the corresponding result R_i is associated. The association is performed by selecting the corresponding granules from $Rep(R_i)$ to the granules linked in $\mathbf{R_{IS-A}}$. Then, R_i is recursively associated with the parent granules of g_k in the hierarchy, and $Ass_{\mathbf{G_O}}(g_k)$ is the set containing such granules.

Single-label Categorization. The identification of the more relevant topical granule is chosen by analyzing the set $Rep(R_i)$ of each search result obtained as described in the previous paragraph, i.e. Multi-label categorization. Now, the selection of the granule is made by considering how the granules in $Rep(R_i)$ are organized in $\mathbf{R_{IS-A}}$; then, the granule is chosen by considering the common parent granule.

A Simple Example. Let us consider the vocabulary of the following granular view related to the Wine Ontology[2]. The set of granules is $\mathcal{G} := \{$ *Marietta Zinfandel, Mountadan Pinot Noir, Lane Tanner Pinot*

[2] http://www.w3.org/2001/sw/WebOnt/guide-src/wine

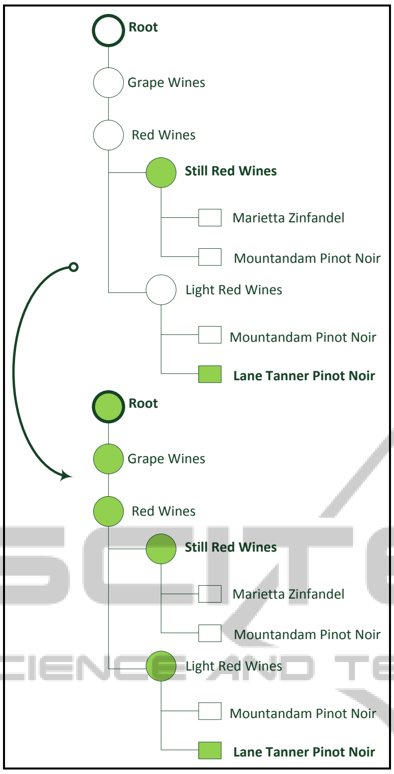

Figure 1: Multi-label categorization of the search result R_1 provided by the granular view.

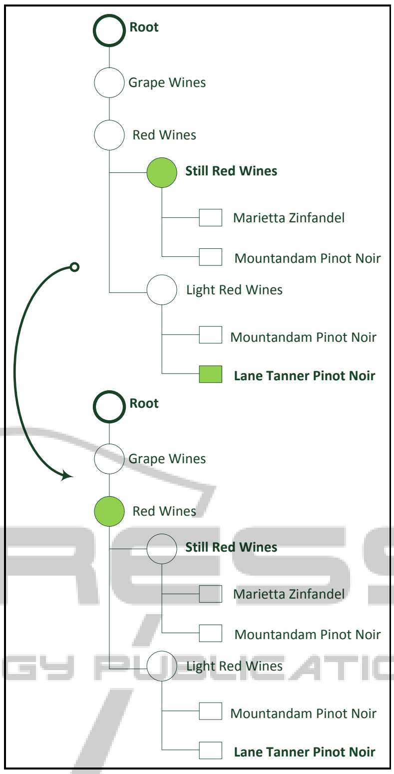

Figure 2: Single-label categorization of the search result R_1 provided by the granular view.

Noir, Red Wines, Grape Wines, Still Red Wines, Light Red Wines}. During a search session a user is interested in finding, for instance, information about *red wines* and he/she writes the following query *q*=*"red wines in France"*, and a list of search results is displayed.

By analysing the first result R_1, we have: *Title="A guide to French Still Red Wines"* and *Snippet="Discover the wines of France, their history an d valleys;... Lane Tanner Pinot Noir is a famous red wine produced in... "*. From these two texts, we index R_1 by the granules of $\mathbf{G_O}$. We obtain $Title(R_1)$ =*Still Red Wine* and $Snippet(R_1)$ =*Lane Tanner Pinot Noir*. Thus, $Rep(R_1) = Title(R_1) \cup Snippet(R_1)$={*Lane Tanner Pinot Noir,Still Red Wines*}.

By considering the case of multi-label categorization, Figure 1 depicts the situation after the application of Step 1 and Step 2 where the result R_1 has been categorized in the following granules *Lane Tanner Pinot Noir, Red Wines, Grape Wines, Still Red Wines, Light Red Wines*. The procedure is repeated for each search result in *Res*.

By considering the case of single-label categorization for $Rep(R_1)$, Figure 2 shows that the single chosen granule is *Red Wines* as it is the common parent granule of *Lane Tanner Pinot Noir* and *Still Red Wines*. Again, the process is repeated for each search result in *Res*.

At the end, the user gets the search results organized as a tree of granules according to the considered categorization methodology, i.e. multi-label or single-label, respectively.

3 EVALUATIONS

As previously outlined, the first objective is to compare the effectiveness of the approach proposed in (Calegari et al., 2011) for the task of single-label categorization with the effectiveness of the cited single-label classifiers in the Introduction as defined in the Weka[3] toolkit. To this aim the standard IR measures of agreement, recall and precision are adopted. The second objective is to evaluate the considered method for the multi-label categorization task; in (Calegari et al., 2011) a comparative evaluation of the multi-label categorization performed on the original domain ontology with respect to the multi-label categorization on its granular view has been presented in terms of agreement, recall and precision. In this paper we

[3] http://www.cs.waikato.ac.nz/ml/weka/

extend and report comparative evaluations based on different indicators, i.e. the NDCG, and the S-Recall measures.

DCG, *discounted cumulative gain*, measures the gain of an item based on its position in the result list. In our context an item refers to a Web search result, and for our experiments we adopted the modified NDCG formulation proposed in (Agrawal et al., 2009). This modification explicitly models a judgment value in addition to the ranking obtained by the application of the methodology, and it normalizes the DCG values by comparing them to an ideal rank given by domain experts. In our case, the judgment assigned by domain experts to each Web search result can be either 0 (Bad) or 1 (Excellent). Formally, given a query q:

$$NDCG(q) = \left(\sum_{g_k \in \mathcal{G}_q} \frac{DCG(Ass_{\mathbf{G_O}}(g_k))}{DCG(Ass_{exp}(g_k))} \right) / |\mathcal{G}_q|,$$

where $\mathcal{G}_q = \bigcup_{R_i \in Res} Rep(R_i)$, and $Ass_{exp}(g_k)$ is the association performed by experts for each search result.

The S-Recall defined in (Zhai et al., 2003) is based on "arguments"; in our context arguments are both topics and subtopics. S-Recall evaluates the size of the set containing the granules obtained in response to a query q over the total number of granules $\mathcal{G}$ as

$$\text{S-Recall}(q) = \frac{|\bigcup_{R_i \in Res_q} C_\alpha(R_i)|}{|\mathcal{G}|},$$

where $C_\alpha(R_i)$ is the set of granules with at least an associated result produced by the categorization process.

4 EXPERIMENT

The method has been implemented as a standalone service that interacts with the Yahoo! Search Engine, and which returns to the user the categorized search results. We conducted the experiments by using the Wine Ontology defined by the Stanford University[4]. A granular view has been obtained from the Wine ontology by following the methodology in (Calegari and Ciucci, 2010). The first experiments aim to analyze the behaviour of the considered categorization approach with respect to the standard single-label classifiers mentioned in Section1. To this aim, with each search result only the broader granule (in the hierarchy) is associated by analyzing the set of granules $Rep(R_i)$.

To generate the domain dependent data set related to the Wine domain, we asked to four wine experts to define 44 topical queries and next, 880 search results have been selected. To this set of results we applied the standard test split with a percentage 70-30 to select training and test set, respectively. In order to approximate a uniform distribution to improve the classifiers prediction capability, we applied a random re-sample on the training set.

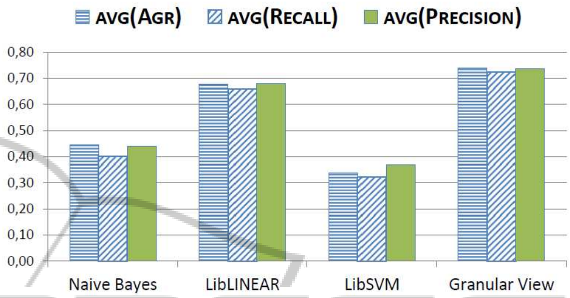

Figure 3: Single-label Categorization.

To compare the considered categorization method (Calegari et al., 2011) with the single-label categorizations on the same data set, during the training phase, the wine experts have acted as assessors in order to labeling each search result with the more appropriate granule of the granular view of the Wine Ontology. Then, the training phase has been applied on 1814 instances. Each instance is characterized by the following attributes: the query, the search result, and the granules set by the domain experts. For the test phase, it has been necessary to set some parameters for the LibSVM and the LibLIN classifiers. LibSVM has been run with the radial basis kernel function with $\nu = 1.0$, while for LibLIN the *cost&bias* parameter had a unitary value with the L_2 metric. The test phase has been performed on 687 instances and for each instance the granule chosen by the considered approaches (i.e., the Multinomial Naïve Bayes, LibLinear, LibSVM, and the single-label categorization of Section 2, named *Granular View*) has been compared with the one identified by the experts by defining the optimal solution.

Figure 3 shows that the *Granular View* method improves the considered standard single-label classifiers. It happens because the standard classifiers do not consider the semantic that each search result can have. In particular, when a search result can semantically belong to more than one granule, the problem is to identify the granule that better represents the several meanings. The single-label categorization approach explained in Section 2 shows close results with respect to the expert expectation by preferring the common coarsest granule among the selected ones. This way, the multiple meaning of a search result are obtained more precisely.

[4] http://www.w3.org/2001/sw/WebOnt/guide-src/wine

To this experiment, the Multinomial Naïve Bayes has been chosen as it is one of the most adopted classifiers and it is based on the assumption that each attribute/feature is independent with respect to the other ones. This is the case suited for our evaluations as each search results has to be classified with a granule and its choice is independent with respect to the other granules used in the attributes set. Support Vector Machine and linear classifiers are de-facto standards for classification problems. Even with a limited population, LibLinear and LibSVM classifier can identify the optimal solution by obtaining results slightly less accurate than what obtained with our single-label categorization approach.

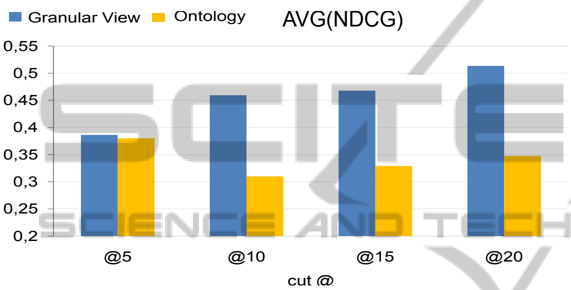

Figure 4: NDCG: Multi-label categorization.

By the second experiment, the method proposed in (Calegari et al., 2011) has been applied to compare the hierarchical multi-label categorization based on the original domain ontology (by only considering the subsumption relation) versus the hierarchical multi-label categorization based on the granular view of the original domain ontology.

The granular view is made up of 162 nodes, whereas the taxonomy of the domain ontology is made up of 219 nodes. We have compared the categorization on a set of 44 queries defined by four wine experts, and we analyzed the top 20 results at different precision levels: @5, @10, @15 and @20. *NDCG*@5 indicates that both taxonomies consider the same results with the same ranking order, and then the measure discounts them in a similar way. The lower $NDCG(q)$ values indicate that the taxonomy of the domain ontology gives to the results a different categorization with respect to the ones assigned by experts; instead, the use of the granular view produces higher $NDCG(q)$ as it preserves the ranking.

The S-Recall(q) evaluations for the categorization provided by experts and both the taxonomies have highlighted that a very small portion of information is involved for each query. Table 1 shows that the domain ontology exhibits a worse behaviour than its granular view: several unnecessary nodes are in fact used to categorize the search results with the consequence that the user spends a lot of time in navigating the hierarchy for discovering the search results.

Table 1: S-recall: multi-label categorization.

cut@	Expert	Granular View	Domain Ontology
@5	0.02	0.10	0.22
@10	0.03	0.15	0.28
@15	0.04	0.19	0.21
@20	0.05	0.34	0.41

5 CONCLUSIONS

In this paper we have presented some evaluations of a method aimed at categorizing search results based on a granular view of a domain ontology to both multi-label categorization and single-label categorization. In this last case, we have proposed an extension of the considered multi-label categorization method in order to manage cases of single-label categorization on search results. This way, it is possible to compare the method with standard classifiers such as the Multinomial Naïve Bayes, the LibLinear, and the LibSVM, respectively.

To test the effectiveness of the considered multi-label categorization approach, we have performed new evaluations in addition to (Calegari et al., 2011) such as NDCG and S-Recall. The granular view has exhibited a better behaviour than the original domain ontology in multi-label categorization, and it has achieved good results if compared to standard single-label classifiers.

As a future research, we will consider other domains, and we will evaluate the approach compared to other standard hierarchical and deep classifiers (e.g., C4.5, random forest, Boltzman machines).

REFERENCES

Agrawal, R., Gollapudi, S., Halverson, A., and Ieong, S. (2009). Diversifying search results. In *WSDM '09*, pages 5–14, New York, NY, USA. ACM.

Baeza-Yates, R. A. and Maarek, Y. (2011). Web retrieval: the role of users. In Ma, W.-Y., Nie, J.-Y., Baeza-Yates, R. A., Chua, T.-S., and Croft, W. B., editors, *SIGIR*, pages 1303–1304. ACM.

Calegari, S. and Ciucci, D. (2010). Granular computing applied to ontologies. *International Journal of Approximate Reasoning*, 51:391–409.

Calegari, S., Farina, F., and Pasi, G. (2011). Topical categorization of search results based on a domain ontology. In Amati, G. and Crestani, F., editors, *ICTIR*, volume 6931 of *Lecture Notes in Computer Science*, pages 262–273. Springer.

Calegari, S. and Pasi, G. (2008). Personalized ontology-based query expansion. In *Web Intelligence/IAT Workshops*, pages 256–259. IEEE.

Calegari, S. and Pasi, G. (2010). Ontology-based information behaviour to improve web search. *Future Internet*, 2(4):533–558.

Carpineto, C., Osinski, S., Romano, G., and Weiss, D. (2009). A survey of web clustering engines. *ACM Computing Surveys*, 41(3):17:1–17:38.

Daoud, M., Tamine-Lechani, L., and Boughanem, M. (2008). Using a graph-based ontological user profile for personalizing search. In *CIKM '08: Proceeding of the 17th ACM conference on Information and knowledge management*, pages 1495–1496, New York, NY, USA. ACM.

Halpin, H. and Lavrenko, V. (2011). Relevance feedback between hypertext and semantic web search: Frameworks and evaluation. *Web Semant.*, 9(4):474–489.

Lv, Y. and Zhai, C. (2010). Positional relevance model for pseudo-relevance feedback. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '10, pages 579–586, New York, NY, USA. ACM.

Ma, Z., Pant, G., and Sheng, O. R. L. (2007). Interest-based personalized search. *ACM Trans. Inf. Syst.*, 25.

Ren, A., Du, X., and Wang, P. (2009). Ontology-based categorization of web search results using YAGO. In *In Comput. Sciences and Optimiz.*, pages 800–804. IEEE.

Ruthven, I. and Lalmas, M. (2003). A survey on the use of relevance feedback for information access systems. *Knowl. Eng. Rev.*, 18(2):95–145.

Salton, G. and Buckley, C. (1990). Improving retrieval performance by relevance feedback. *Journal of the American Society for Information Science*, 41:288–297.

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Comput. Surv.*, 34(1):1–47.

Shahabi, C. and Chen, Y.-S. (2003). Web information personalization: Challenges and approaches. *DNIS. LNCS.*, 2822:5–15.

Teevan, J., Dumais, S. T., and Horvitz, E. (2005). Personalizing search via automated analysis of interests and activities. In Baeza-Yates, R. A., Ziviani, N., Marchionini, G., Moffat, A., and Tait, J., editors, *SIGIR*, pages 449–456. ACM.

Zhai, C., Cohen, W. W., and Lafferty, J. (2003). Beyond independent relevance: Methods and evaluation metrics for subtopic retrieval. In *In Proceedings of SIGIR*, pages 10–17.