

# Multi-Dimensional Reputation Modeling using Micro-Blog contents

Jean-Valère Cossu, Eric SanJuan(✉), Juan-Manuel Torres-Moreno  
and Marc El-Bèze

LIA/Université d'Avignon et des Pays de Vaucluse  
339 chemin des Meinajaries, Agroparc BP 91228, 84911 Avignon cedex 9, France  
`firstname.name@univ-avignon.fr`<sup>1</sup>

**Abstract.** In this paper, we investigate the issue of modeling corporate entities' online reputation. We introduce a bayesian latent probabilistic model approach for e-Reputation analysis based on Dimensions (Reputational Concepts) Categorization and Opinion Mining from textual content. Dimensions to analyze e-Reputation are set up by analyst as latent variables. Machine Learning (ML) Natural Language Processing (NLP) approaches are used to label large sets of text passages. For each Dimension, several estimations of the relationship with each text passage are computed as well as Opinion and Priority. The proposed automatic path modeling algorithm explains Opinion or Priority scores based on selected Dimensions. Model Robustness' is evaluated over RepLab dataset.

## 1 Introduction

Online Reputation Management is recently being a top-notch subject of studies investigated in several researches from various fields such Marketing, Psychology, Social Media Analysis (SMA), Information Retrieval (IR) and NLP. Although researchers have been processing and managing business information for several decade, ML techniques are still lacking to support the process of providing a relevant Reputation Modeling. An increasing number of SMA services and community managers use Micro-Blog streams from Twitter based on key-word queries to analyze Opinion and Reputation related to some entities. The use of key-word queries instead of analyzing the whole Micro-Blog sphere allows to deal with very large Micro-Blog networks and to avoid complex an expensive pooling strategies, often non-accurate for a micro analysis of specific entities like museums in Barcelona at the time of champion league. The drawback of these streams based on key-word queries is that they only give a partial and incomplete view of the whole Micro-Blog network. Removing or adding some keywords can lead to very different datasets. Moreover, Dimensions (Reputational Concepts) to explain reputation and recommend actions are often vague and depend on individual analyst expertise and background that could be misled by a peculiar choice of keywords. Given an entity (public figure), a set of pre-defined attributes, we seek to model their impact on two objective measures:

1. External public Opinion expressed about the entity,
2. Internal Stimuli Priority [1] ,

based on a stream of public Micro-Blogs mentioning the entity.

## 2 Probabilistic Path Modeling

We consider the case where the analyst defines a set of  $D$  Dimensions through which he aims to analyze Opinion or to define Priority over a set of Micro-Blogs. To define these Dimensions we suppose that entity provides examples of Micro-Blogs for each Dimension and that for each example it indicates the Opinion (Positive, Neutral, Negative) and the Priority (Alert, Important, Non-Important). The chosen Dimensions were suggested by e-watcher specialists from the Reputation Institute’s Reprtrak like: Products and services, Citizenship, Governance, Innovation, Leadership, Performance, Workplace.

Therefore, given a trend  $\Omega$  of Micro-Blogs texts, the analyst provides a family  $\mathcal{S} = S_1, \dots, S_D$  of subsets of  $\Omega$ , each representing a Dimension, an Opinion or an "Alert" that we shall model as a latent variable. We do not expect  $S_i$  to be unbiased pools of Micro-Blogs, but we can still use them to train a sequence of  $K$  independent classifiers over each set that will evaluate the possibility for a new Micro-Blog text to be included by the analyst in some  $S_i$ . More formally we consider  $K$  sets of  $D$  scoring normalized functions such as:

$$f_{i,k} : \omega \in \Omega \mapsto f(\omega) \in [0, 1] \quad (1)$$

$$\omega \in S_i \Rightarrow (\forall 1 \leq i \leq D) f_{i,k}(\omega) = 1 \quad (2)$$

Each normalized scoring function  $f_{i,k}$  defines a discrete smoothed probability function  $P_{i,k}$  over  $\Omega$  defined by:

$$P_{i,k}(\{\omega\}) = (1 - \lambda) f_{i,k}(\omega) + \lambda E(f_{i,k}) \quad (3)$$

where  $\lambda$  is a smoothing parameter in  $[0, 1]$  and  $E(f_{i,k})$  is the expectation of  $f_{i,k}$  over a trend of Micro-Blogs and can be simply estimated over large finite pools  $T$  of Micro-Blogs as:  $\frac{1}{|T|} \sum_{\omega \in T} f_{i,k}(\omega)$ . We hypothesize that for each latent variable  $S_i$  there exists:  $\alpha_i = (\alpha_{i,1}, \dots, \alpha_{i,k})$ ,  $\alpha_{i,k} > 0$  such that the real probability  $P_i$  for a Micro-Blogs  $\omega$  to be associated by the analyst with  $S_i$  verifies:

$$P_i(\{\omega\}) = \prod_{k=1}^K P_{i,k}(\{\omega\})^{\alpha_{i,k}} \quad (4)$$

To analyze Opinion and Priority based on entity’s defined Dimensions, we define two path models among latent variables. For Opinion, let us suppose that  $S_D$  is associated with "Neutral",  $S_{D-1}$  with "Negative" and  $S_{D-2}$  with "Neutral" meanwhile  $S_1, \dots, S_{D-3}$  are the entity’s defined Dimensions. Then we consider as path model the acyclic directed graph  $G_{op}(\mathcal{S}) = (1, \dots, D, A_{op}(\mathcal{S}))$  where the set the set of arrows  $A_{op}(\mathcal{S})$  is defined as:

$$A_{op}(\mathcal{S} = \{1, \dots, D-3\} \times \{D-2, D-1\} \cup \{D-2, D-1\} \times \{D\}) \quad (5)$$

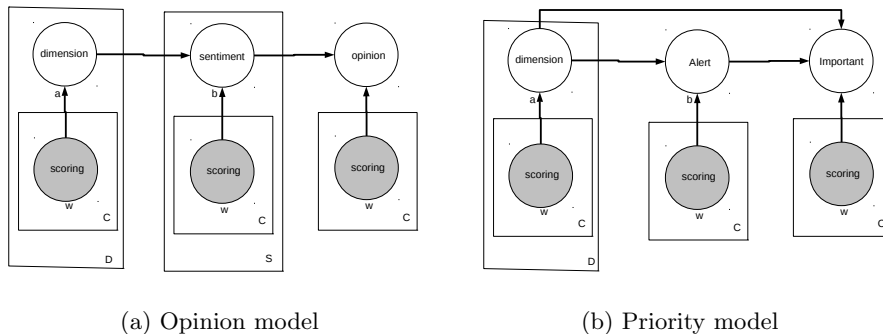


Fig. 1: Plate notation for path model

Figure 1a gives the plate representation of this path model. For Priority we define a slightly different path model: we suppose that  $S_D$  is associated with “Important” and  $S_{D-1}$  with “Alert” meanwhile  $S_1, \dots, S_{D-2}$  are the entity’s defined Dimensions. Then we consider as path model the acyclic directed graph  $G_{al}(\mathcal{S}) = (1, \dots, D, A_{al}(\mathcal{S}))$  where the set the set of arrows  $A_{al}(\mathcal{S})$  is defined as:

$$A_{op}(\mathcal{S} = \{1, \dots, D-2\} \times \{D-1, D\} \cup \{(D-1, D)\}) \quad (6)$$

Figure 1b gives the plate representation of this path model. To analyze the impact of Dimensions over Opinion or “Alert”, given a smoothed parameter  $\lambda$ , we propose to seek for the probabilistic model with parameter  $\alpha = [\alpha_{i,k}]_{1 \leq i \leq D, 1 \leq k \leq K}$  that maximizes the product of conditional probabilities along the Bayesian network induced by  $G \in \{G_{op}, G_{al}\}$ , i.e.:

$$\alpha(\lambda, \mathcal{S}) = \underset{\alpha}{\operatorname{argmax}} \left\{ \prod_{(i,j) \in G} P(S_i/S_j) \right\} \quad (7)$$

### 3 Experimental data and settings

#### 3.1 Replab Dataset

We use the context of RepLab [2, 3] tasks to evaluate our proposal that is to say: to propose an overview of 61 entity’s (drawn in 4 domains: Automotive, Banking, Music and University) e-Reputation regarding experts taxonomies using provided set and pertaining of Micro-Blogs concerning each entity. Data

have been annotated by experts from Llorente & Cuenca<sup>1</sup> for either Polarity for Reputation (3 levels), Priority (3 levels) and Reputation Dimensions (only for Automotive and Banking domains). These Dimensions have been identified as relevant interests and are considered as key issues for company' stakeholders.

### 3.2 Scoring Protocol

**Term Scoring** The proposed supervised corpus-based classification method is entirely based on the tweets contents, i.e., it does not use any external knowledge resource. Besides being the most relevant source of information for classification, the text cannot be directly processed by classifiers and learning algorithms in its unstructured natural form. For that reason, our tweets should be translated into a manageable form: the discriminant bag-of-words representation. In this step, to each term multiple weights (using Frequency-Inverse Document Frequency (TF-IDF-Gini) approach [12]) are assigned, which describe the strength of association of a term with each label, Priority and Polarity levels. We assign a score to each term corresponding to its strength of association with each Dimensions, Priority and Polarity (borrowed from the tweet's label in the training set). Terms mentioned frequently in tweets that have a specific label gets the highest score.

**Document Scoring** For an evaluation purpose RepLabs' organizers provided a single annotation for each tweet which correspond to discrete binary score and may remain insufficient to provide a reputation modeling. Dimensions like Innovation or Leadership are vague. Moreover tweets can be misleading and ambiguous this is why we undertake a scoring approach rather than classification using strong learning references. In this step, using statistical NLP approaches [12], we project each document in a #Dimensions  $\times$  #Opinion or Priority  $\times$  #Classifiers dimensions spaces. We consider the similarities between each tweet and each label using: Cosine distance, Jaccard index, Linear SVM and k-NN.

**Scoring Normalization** As we aim at ranking document in each class, all similarities need to be normalized to the same scale. For each tweet, the degree of similarity of similarity (confidence degree) to each class provided by kNN, Cosinus and Jaccard have been normalized using the sum of all system-provided scores related to this document. For SVM we considered the lowest hypothesis score as a rejected class. Its value was added to all other classes, then it has been normalized just as mentioned before. Cosinus and Jaccard proposals fit the typical IR issue as they rank each document in each class while kNN and SVM approach are nearer human annotator behavior in evaluating the document with a limited number of background neighbors and rejecting hypothesis. The similarity measure operate only on the overlapping document and some documents are assigned with a zero weight which limit the abstraction process and increase the probability of strong errors. It leads us to consider probability re-estimation of a document  $d$  in a class  $c$  using a smoothing as defined in (3).

<sup>1</sup> <http://www.llorenteycuenca.com/>

## 4 Dimension Ranking Evaluation

Rather than looking for classification performances we focus on Reputation Modeling. When considering only the best class hypothesis provided by classifiers, the Document Scoring Approach performances have been reported in [12]. We use Replab testing set to evaluate the ability of our probabilistic path model to rank Dimensions by decreasing impact over an Opinion or a Priority. We expect models presented here to be independent from the selected text classifiers used to score Dimensions and Opinions. We seek for models' robustness against incompleteness of data used as training sets. RepLab complete reference are available for two domains: Automotive and Banking. For Banking the test set contains 10,781 tweets among which 4,524 are Negative, 5,015 Positive, the rest being Neutral. Our probabilistic model provides two rankings of these Dimensions:

1. Impact over Negative opinions scored by  $P(Negative/Dimension)$
2. Impact over Positive opinions scored by  $P(Positive/Dimension)$

We compare each of these rankings against the reference:

$$cor(ref_{Dimension}, ref_{Negative} - ref_{Positive}) \quad (8)$$

For each Dimension  $ref_{Dimension}$  is a binary vector (1 if the tweet is in this Dimension, 0 otherwise) meanwhile  $ref_{Negative} - ref_{Positive}$  is ternary (1 if Negative, 0 if Neutral, -1 if Positive). The results of the ranking produced by  $P(Negative/Dimension)$  are the following:

- Pearson's product-moment correlation = 0.9759958,  $p$ -value = 0.0008574
- Kendall's rank correlation  $\tau = 1$ ,  $p$ -value = 0.002778
- Spearman's rank correlation  $\rho = 1$ ,  $p$ -value = 0.002778

Therefore, for Banking rankings towards Negative are highly correlated and these correlations are statistically highly significant, meanwhile the same tests over the training corpus are not significant because of data incompleteness. The ranking produced by  $P(Positive/Dimension)$  is less correlated because of Neutral Micro-Blogs but correlation remains statistically significant ( $p$ -value < 0.1).

For Automotive, there are 12,776 tweets in the reference with only 1,718 Negative against 10,995 Positive and only 63 Neutral. The higher correlation is obtained comparing  $P(Negative/Dimension)$  against :

$$cor(ref_{Dimension}, ref_{Negative}) \quad (9)$$

In this case Spearman's rank correlation  $\rho = 0.7714286$  but  $p$ -value=0.1028. All other correlations are positive but non significant. In Banking and Automotive cases, the probabilistic model performs better on explaining Negative than Positive. In both cases the Negative class is narrow than the Positive class.

Finally, for *Alerts* in the case of Banks, Dimensions are ranked in the exact same order. However, the Pearson's product-moment correlation between the predicted inner latent variable *Alert* and the reference is 0.49 is significantly high ( $p$ -value <  $10^{-3}$ ) but not significantly higher than the single Cosine estimate. This means that in this case, the use of a single basic classifier would have provided similar results which is not the case for Opinion.

## 5 Discussion and conclusions

In this paper we examined an automatic method to order relations between classes and infer more complex latent hierarchies. For that we introduced a probabilistic approach for e-Reputation analysis based on Reputational Concepts Categorization and Opinion Mining from textual content. For each Concepts, several estimations of the relationship with each text passage are computed. We proposed an automatic path modeling algorithm to explain Opinion or Priority scores based on selected Concepts. Robustness of the resulting model has been evaluated over the multilingual CLEF RepLab dataset.

**Acknowledgment** This work is funded by the project ImagiWeb ANR-2012-CORD-002-01.

## References

1. Mather, M., Sutherland, M.R.: Arousal-biased competition in perception and memory. *Perspectives on psychological science* **6**(2) (2011) 114–133
2. Amigó, E., De Albornoz, J.C., Chugur, I., Corujo, A., Gonzalo, J., Martín, T., Meij, E., De Rijke, M., Spina, D.: Overview of replab 2013: Evaluating online reputation monitoring systems. In: CLEF. (2013)
3. Amigó, E., Carrillo-de Albornoz, J., Chugur, I., Corujo, A., Gonzalo, J., Meij, E., de Rijke, M., Spina, D.: Overview of replab 2014: author profiling and reputation dimensions for online reputation management. In: CLEF. (2014)
4. Villena Román, J., Lana Serrano, S., Martínez Cámara, E., González Cristóbal, J.C.: Tass-workshop on sentiment analysis at sepln. (2013)
5. Zhao, W.X., Jiang, J., He, J., Song, Y., Achananuparp, P., Lim, E.P., Li, X.: Topical keyphrase extraction from twitter. In: Proceedings of the 49th Annual Meeting of the ACL: Human Language Technologies. (2011)
6. Velcin, J., Kim, Y., Brun, C., Dormagen, J., SanJuan, E., Khouas, L., Peradotto, A., Bonnevey, S., Roux, C., Boyadjian, J., et al.: Investigating the image of entities in social media: Dataset design and first results. In: LREC. (2014)
7. Peleja, F., Santos, J., Magalhães, J.: Reputation analysis with a ranked sentiment-lexicon. In: Proceedings of the 37th SIGIR conference. (2014)
8. McDonald, G., Deveaud, R., McCreddie, R., Macdonald, C., Ounis, I.: Tweet enrichment for effective dimensions classification in online reputation management. In: Ninth International AAAI Conference on Web and Social Media. (2015)
9. Qureshi, M.A., O’Riordan, C., Pasi, G.: Exploiting wikipedia for entity name disambiguation in tweets. In: NLP and Information Systems. (2014)
10. Derczynski, L., Maynard, D., Rizzo, G., van Erp, M., Gorrell, G., Troncy, R., Petrak, J., Bontcheva, K.: Analysis of named entity recognition and linking for tweets. *Information Processing & Management* (2015)
11. Damak, F., Pinel-Sauvagnat, K., Boughanem, M., Cabanac, G.: Effectiveness of state-of-the-art features for microblog search. In: The 28th ACM Symposium on Applied Computing. (2013)
12. Cossu, J.V., Janod, K., Ferreira, E., Gaillard, J., El-Bèze, M.: Nlp-based classifiers to generalize experts assessments in e-reputation. In: CLEF. (2015)