

# Automatic Classification and PLS-PM Modeling for Profiling Reputation of Corporate Entities on Twitter

Jean-Valère Cossu<sup>(✉)</sup>, Eric Sanjuan, Juan-Manuel Torres-Moreno,  
and Marc El-Bèze

LIA/Université d'Avignon et des Pays de Vaucluse, 39 chemin des Meinajaries,  
Agroparc BP 91228, 84911 Avignon cedex 9, France  
[jean-valere.cossu@univ-avignon.fr](mailto:jean-valere.cossu@univ-avignon.fr)

**Abstract.** In this paper, we address the task of detecting the reputation alert in social media updates, that is, deciding whether a new-coming content has strong and immediate implications for the reputation of a given entity. This content is also submitted to a standard typology of reputation dimensions that consists in a broad classification of the aspects of an under public audience company. Reputation manager needs a real-time database and method to report what is happening right now to his brand. However, typical Natural Language Processing (NLP) approaches to these tasks require external resources and show non-relational modeling. We propose a fast supervised approach for extracting textual features, which we use to train simple statistical reputation classifiers. These classifiers outputs are used in a Partial Least Squares Path Modeling (PLS-PM) system to model the reputation. Experiments on the RepLab 2013 and 2014 collections show that our approaches perform as well as the state-of-the-art more complex methods.

## 1 Introduction

Recently, with the emerging trend of the online networked information, control has moved to users. Each act of a public entity become scrutinized by a powerful global audience. We can understand reputation as the general recognition by other people of some characteristics for a given entity. Specifically, in business or politics, reputation comprises the decisions taken and how it is perceived by the population. This requires new reputation management tools and strategies able to consider the variability of interpretations for a given document. The rise of online social media has become an interesting way to process large amount of opinions about entities even though in the case of tweets there are no explicit ratings to be directly used in an opinion processing. Although significant advances have been made in RepLab<sup>1</sup> [1, 2]. Analyzing reputation about companies and individuals is a hard problem requiring a complex modeling of these entities and

---

<sup>1</sup> Replab provides a framework to evaluate Online Reputation Management systems on Twitter <http://www.limosine-project.eu/events/replab2013>.

it is still a significant research challenge because unlike products, opinions about entities may vary from the point-of-view (the person who speaks, who reads) and the context. This modeling can be considered similar as tweets taxonomy or and handmade classification.

In this paper, we are interested in reputation alert detection and reputation dimension assignment. It consists in identifying relevant interests for stakeholders in the companies that are also considered as key issues for the entities reputation, in order to contribute to a better understanding of a group of tweets and their topic. These aspects are subjective and may depend on each expert. We aim at guiding this expert to consider that a given tweet remains more important than another one. We use NLP based classifiers to project each tweet in the classes space in order to generalize the experts point-of-view. Then by observing interactions between classes using PLS-PM [3] we can provide a hierarchy to visualize how tweets and classes are perceived in the tweets stream.

## 2 Related Work

Previous research has exploited supervised methods for topic categorization of short social chat messages as [4]. But these kind of required a costly human annotations which is usually not available in real time for large-scale micro-blogging messages. Due to a lack of applicable performance metrics and exploitable gold-standard labels, it is hard to report the systems performance, or comment on the generalization of approaches such as TweetMotif [5]. Most of the contributions on reputation monitoring to extract sets of tweets requiring a particular attention from a reputation manager have been proposed in the last editions of RepLab [1,2]. Issues were tackled with Social Network Analysis [6] and both supervised and unsupervised algorithms [7] combined with terms selection strategies. Joint work between tweet clustering and priority detection over a NLP-based classification was also proposed by [8]. RepLab'2014 [2] focused on the reputation dimension classification. Which can be viewed as a complement to topic detection it is nearer a stress classification of the company's aspects. The stresses are introduced by the experts and only reflect their interest. Some approaches considered information beyond the tweet textual content such as pseudo-relevant document [9] or semantic expansion [10] and Wikipedia categories [11] but also psychometric and linguistic information [12].

In spite of the great significance of extracting information to obtain high performances, the amount of research dedicated to understand the experts' stress effect is really limited. This leads us to investigate not the best possible performance but to propose a toolbox that allows a better understanding of scores given by classifiers for each tweet-content.

## 3 Outer Model Learning

We understand the problem of detecting reputation using a supervised classification method based on a threshold intersection graph computed over the discriminant bag-of-words representation of each tweet. Vertices are tweets, edges

are between tweets sharing at least one word (lexical relation) and are weighted using Cosine distance and Jaccard similarity index which we add Multi-Class Support Vector Machin<sup>2</sup>. We start with annotated documents based on their following reputation tags: priority level and dimension. We use Term Frequency-Inverse Document Frequency (TF-IDF) [14, 15] combined with the Gini purity criterion [16]. Then we estimate the similarity of a given tweet by comparing it to each class bag-of-words and rank tweets according to the score provided by the classifiers.

## 4 Inner Model Learning

### 4.1 PLS-PM Models

The statistical method PLS-PM allows to estimate parameters of linear regression by calculating the solution of the general underlying model of multivariate PLS [17]. Dealing with ranked classes and several ranking led to meaningfully combine these sets and PLS-PM is an interesting way to combine systems outputs. Each tweet is analyzed like a structure made of blocks of manifest (observed) variables (scores in each class). Each block is summarized by a latent variable. PLS-PM is used to find the best weight (for system-class pairs) to predict a conditional “ALERT” priority level using dimension and priority probabilities given by each classifier.

### 4.2 Model for Profiling Reputation

We propose a model combining pre-defined dimensions suggested by e-watcher specialists<sup>3</sup> like Products and services, Citizenship, Governance, Innovation, Leadership, Performance, Workplace with the concepts of Alert and Importance for Profiling reputation. The objective is to be able to explain why some action should be taken after some tweets based on these dimensions. The model also allows to follow the impact of these dimensions over the time and improves the robustness of alerts.

Each dimension is modeled as a latent variable combining several textual classifiers and entailing one and only one of the two concepts Alert or Importance. Alert is also supposed to entail Importance. Therefore dimensions are separated into 2 groups: those that will directly induce Alert and those that seem to be less strategic (thus considered as unimportant). Since all classifier outputs have been normalized between 0 and 1, the complement of the dimension  $1 - x$  can be considered and supposed to entail Importance: the underlying 3 valued logic (Alert, Important, NonImportant) is that the complement of a *Non Alert* entails *Important*.

---

<sup>2</sup> See [13] [http://www.cs.cornell.edu/people/tj/svm\\_light/svm\\_multiclass.html](http://www.cs.cornell.edu/people/tj/svm_light/svm_multiclass.html).

<sup>3</sup> Reputation Institute’s Reprtrak framework <http://www.reputationinstitute.com/about-reputation-institute/the-reprtrak-framework>.

Dimensions are split according to the learning corpus keeping in mind that a tweet can have multiple dimensions and that annotators agreement cannot be expected to be high since most of these dimensions like Innovation or Leadership are vague. Moreover tweets can be misleading and ambiguous. Alert is also estimated based on these classifiers meanwhile Importance is estimated as non-unimportant. The PLS-PM algorithm is then used to estimate each inner variable as a vector minimizing square distance to both classifiers normalized output scores and related inner latent variables. We use the  $R$  index to estimate the model quality (maximizes the square sum of correlations inside latent variables and between related variables).

## 5 Experimental Evaluation

### 5.1 Evaluation Protocol

We perform a supervised classification over Replab’2013-14 dataset [1,2]. The dataset includes a training set and a test set both annotated in either of several reputation monitoring axis: filtering, polarity, priority-level (*Alert*, *Important*, *Unimportant*), clustering and dimension (only for Automotive and Banking domains). This paper only addresses priority-level and dimension classification issues. We consider lower-cased and cleaned text<sup>4</sup>. We choose to be entity and language independent, since we want to detect a tweet reputation regardless of its association with an entity. We compare our system to RepLab’2013-4 baselines and best systems according to official metrics: Reliability/Sensitivity [18] and the overall accuracy which we added typical F-Score (based on Precision and Recall).

### 5.2 Outer Model Evaluation

**Priority Detection.** The best F-measure (Best F in the tables) reported in this task was obtained with a kNN based classification method [19]. Other performances ranked with regards to F-measure are summarized in Table 1. Both SVM and Cosinus approaches are competitive according to Accuracy. Although SVM

**Table 1.** Priority detection performances ordered by F-Measure (R,S).

System	F-Score	Accuracy	Reliability	Sensitivity	F-Measure (R,S)
<i>Best F</i>	<b>.571</b>	.636	.387	.315	<b>.335</b>
SVM	.553	<b>.643</b>	.344	.294	.304
<i>Baseline</i>	.512	.570	<b>.403</b>	.248	.274
Cosinus	.566	.637	.344	.236	.260
Jaccard	.492	.561	.342	.212	.233

<sup>4</sup> We remove links, stop-words and punctuation marks.

**Table 2.** Dimensions classification task Performances ordered according to Accuracy.

System	F-Score	Accuracy
Cosinus	<b>.505</b>	<b>.741</b>
SVM	.467	.733
<i>Best_Acc</i>	.473	.731
<i>Baseline</i>	.380	.622
Jaccard	.378	.476

shows significant improvement with Accuracy it falls with F-Score in contrast with the Cosinus that completely collapses regarding F(R,S).

**Dimensions Classification.** The best RepLab’2014 participating systems (noted *Best\_Acc* in Table 2) system used tweet enrichment via pseudo-relevant document [9]. As interesting results, we can see that the Cosinus classifier outperforms all performances reported up to now in this task.

### 5.3 Inner Model Evaluation

Based on the Replab Learning data it appeared that there is a common model for all four domains Bank, Automotive, Music and University in this reputation task: three dimensions entail Alert: Governance, Innovation, Leadership, Performance and WorkPlace. The complement of the two remaining dimensions entail Importance (i.e.: NonUnimportant): NonCitizenship and NonProduct & Services. Figure 1 shows the inner model with its estimated coefficients using the R PLS-PM library<sup>5</sup> where all variables are estimated using Cosine, Jaccard and SVM classifier scores. Figure 2 shows the same model but where inner variable Alert has been replaced by the reference from CLEF Replab. The standard deviation between the predicted model and the reference is significant (t-test p-value < 0.05) but the dimensions are ranked in the exact same order. The Pearson’s product-moment correlation between the predicted inner latent variable Alert and the reference is 0.49 which is significantly high (p-value < 10<sup>-3</sup>) but not significantly higher than the single Cosine estimate. Therefore, for the bank domain, the PLS-PM model helps in prioritizing dimensions but not in improving alert prediction. For the automotive domain the results are slightly different, the ranking of dimensions is not exactly the same since the predictive model ranks Governance (0.53 path coefficient towards Alert) before Innovation (0.15) meanwhile in the reference Innovation has an higher impact on Alert (0.23 vs 0.14). However both rankings are highly correlated (Kendall test: 80 <  $\tau$  < 1, p-value < 0.05) and the inner variable Alert is significantly highly correlated to the reference (Pearson 95 percent confidence interval between 0.24 and 0.25). For

<sup>5</sup> <http://cran.r-project.org/web/packages/plspm/>.

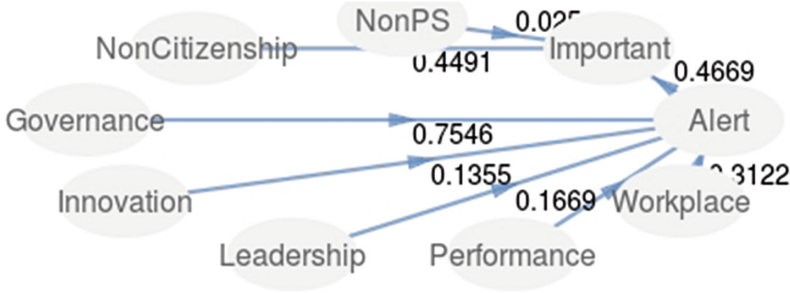


Fig. 1. Inner Model for Bank domain with inferred Alert

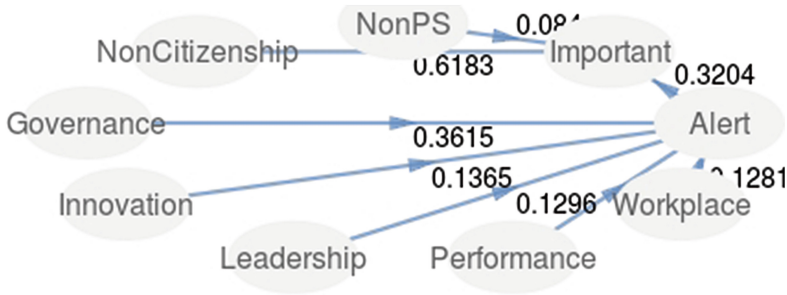


Fig. 2. Inner Model for Bank domain with reference Alert

the Music and the University domain, correlations are lower (Pearson’s product-moment correlation of 0.15 and 0.11) because classifiers were less efficient on these domains for Alert prediction, but PLS-PM models significantly improves them by 10%. Even though we do not have a reference annotation for dimensions in these domains. Classifiers have been trained without domain distinction, since Bank and Automotive are the largest domains, they tend to expand rules from Banks to non related domains but do not infer inverted relations.

## 6 Conclusions

The experimental evaluations on RepLab establish that tweets lexical content is sufficient to tackle the tasks of identifying the reputation alerts and dimensions of micro-blog posts using simple machine learning approaches. We experimented simple and more complex statistical lexical corpus-based NLP methods that use discriminating textual features inferred from labeled data. Our approaches turn out to be very effective in addressing the priority and reputation dimensions detection task. It then appeared that PLS-PM modeling based on a three valued scale: *Alert*, *Important*, *Unimportant* could compensate this lack by detecting uncorrelated tweets with existing topics. In future work, we plan to examine

an automatic method to order relations between classes (dimensions) and infer more complex latent hierarchies. We also intend to study an interesting lexical context expansions simulating an active learning over non-annotated tweets.

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

## References

1. Amigó, E., Carrillo de Albornoz, J., 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: Forner, P., Müller, H., Paredes, R., Rosso, P., Stein, B. (eds.) CLEF 2013. LNCS, vol. 8138, pp. 333–352. Springer, Heidelberg (2013)
2. 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: Kanoulas, E., Lupu, M., Clough, P., Sanderson, M., Hall, M., Hanbury, A., Toms, E. (eds.) CLEF 2014. LNCS, vol. 8685, pp. 307–322. Springer, Heidelberg (2014)
3. Wold, S., Eriksson, L., Trygg, J., Kettaneh, N.: The pls method-partial least squares projections to latent structures-and its applications in industrial rdp (research, development, and production). Unea University (2004)
4. Ranganath, R., Jurafsky, D., McFarland, D.: It’s not you, it’s me: detecting flirting and its misperception in speed-dates. In: Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, vol. 1, pp. 334–342. Association for Computational Linguistics (2009)
5. O’Connor, B., Krieger, M., Ahn, D.: Tweetmotif: exploratory search and topic summarization for twitter. In: ICWSM (2010)
6. Berrocal, J.L.A., Figuerola, C.G., Rodríguez, Á.Z.: Reina at replab2013 topic detection task: community detection. In: CLEF 2013 Eval. Labs and Workshop Online Working Notes (2013)
7. Sánchez-Sánchez, C., Jiménez-Salazar, H., Luna-Ramirez, W.: Uamclyr at replab2013: monitoring task. In: CLEF 2013 Eval. Labs and Workshop Online Working Notes (2013)
8. Cossu, J.V., Bigot, B., Bonnefoy, L., Senay, G.: Towards the improvement of topic priority assignment using various topic detection methods for e-reputation monitoring on twitter. In: Métais, E., Roche, M., Teisseire, M. (eds.) NLDB 2014. LNCS, vol. 8455, pp. 154–159. Springer, Heidelberg (2014)
9. McDonald, G., Deveaud, R., McCreddie, R., Gollins, T., Macdonald, C., Ounis, I.: University of glasgow terrier team/project abacá at replab 2014: reputation dimensions task (2014)
10. Rahimi, A., Sahlgren, M., Kerren, A., Paradis, C.: The stavicta group report for replab 2014 reputation dimensions task. In: CLEF 2014 Evaluation Labs and Workshop-Working Notes Papers (2014)
11. Qureshi, M.A., Oriordan, C., Pasi, G.: Exploiting wikipedia for entity name disambiguation in tweets. In: Métais, E., Roche, M., Teisseire, M. (eds.) NLDB 2014. LNCS, vol. 8455, pp. 184–195. Springer, Heidelberg (2014)

12. Vilares, D., Hermo, M., Alonso, M.A., Gómez-Rodríguez, C., Vilares, J.: Lys at clef replab 2014: Creating the state of the art in author influence ranking and reputation classification on twitter. In: CLEF, pp. 1468–1478 (2014)
13. Crammer, K., Singer, Y.: On the algorithmic implementation of multiclass kernel-based vector machines. *J. Mach. Learn. Res.* **2**, 265–292 (2002)
14. Salton, G., Buckley, C.: Term-weighting approaches in automatic text retrieval. *Inf. Process. Manag.* **24**(5), 513–523 (1988)
15. Robertson, S.: Understanding inverse document frequency: on theoretical arguments for idf. *J. Documentation* **60**(5), 503–520 (2004)
16. Torres-Moreno, J., El-Beze, M., Bellot, P.: Bechet, opinion detection as a topic classification problem in in textual information access. chap. 9 (2013)
17. Henseler, J.: On the convergence of the partial least squares path modeling algorithm. *Comput. Statistics* **25**(1), 107–120 (2010)
18. Amigó, E., Gonzalo, J., Verdejo, F.: A general evaluation measure for document organization tasks. In: Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 643–652. ACM (2013)
19. Cossu, J., Bigot, B., Bonnefoy, L., Morchid, M., Bost, X., Senay, G., Dufour, R., Bouvier, V., Torres-Moreno, J., El-Beze, M.: Lia@ replab 2013. In: CLEF 2013 Eval. Labs and Workshop Online Working Notes (2013)