Content-based classifiers to generalize expert assessments in E-Reputation

J.V. Cossu
University of Avignon, France

CLEF - Best of the Labs 2014

10 September 2015

Participants

- **J.-V. Cossu**, K. Janod, E. Ferreira, J. Gaillard and M. El-Bèze
- Speech and NLP team

Plan

- Introduction : 3 years of RepLab
- Reputation Monitoring
- Author Profiling
- Conclusion

Get started with Online Reputation on Twitter and RepLab

3 years of RepLab

- RepLab 2012 Monitoring in unknown entity case
- RepLab 2013 Monitoring in known entity scenario
- RepLab 2014 Monitoring and user profiling

Objectives

- Experts and entities concerns
- Learn experts behaviour
- Automatically propagate these assessments
- Justify the hypothesis

Tweet 208660000908914688 from Jazzrockman about Justin Bieber

Justin Bieber es el nombre de una bebida afrodisiaca?
EN: Justin Bieber is the name of an aphrodisiac drink?
Tagged as having a positive impact over Justin Bieber's reputation

RepLab: Reputation monitoring

Profiling Corporate entities on Twitter

- Filter unrelated contents
- Group tweets depending on trends
- Detect aspects oriented opinions
- Find tweets needing immediate attention and react appropriately

2013 and 2014 tweets collection

- Large tweets collection (142,000 tweets, 80% English, 20% Spanish)
- 61 entities from Automotive, Banking, Music, Universities
- Crawling with entity name between June and Decembre 2012
- More than 500,000 manual experts assessments
- Finally a few training data for each entity (600 training examples)

What does it look like ?

Tweet 281209605940449280 from stopbeingfamous about Wells-Fargo

Here we have Wells-Fargo funding Mexican drug cartels, the ATF FBI allowing class 3 weapons across the US-Mexico border, HSBC laundering
Related, Negative, Citizenship, Wachovia Helped Mexican Drug Smuggles, Alert

ORM: First outcomes

Solved issues

- Filtering problem is almost solved
- Topic/Event detection
- Polarity for reputation can be efficiently estimated

Still running

- Detecting concept
- Estimate priority

Focused issues

Priority Ranking

Predict the priority level of conversations:

- Negative messages may impact the priority rank
- Message from "opinions makers" may receive an higher priority
- Classification or Relationship: ALERT > MILDLY > NOT IMPORTANT
- Rank the few Alerts (3%) using the probability of being ALERT
- Official evaluation as clustering of relationships

Dimensions Classification

- Dimensions = Conceptual Reputation Aspects suggested by specialists
- *Citizenship*, *Governance*, *Innovation*, *Leadership*, *Performance*, *Products*, *Workplace*
- Most tweets are about Products/Services (60%)
- Official evaluation using typical *Precision/Recall/Accuracy*

Let's return to an example

Tweet 213250100691091456 from TibidyUS about Volvo

Volvo produces first external airbag to protect pedestrians: http://bit.ly/LEAda #Swedish
Related, Positive, Innovation, Pedestrian airbag, Important

- Innovation is rather vague, Leadership category could also fit
- Topic could be Volvo's innovations

Content-based features

- Bag of N-grams representation (TF, IDF, purity index)
- Tweet length, Special characters, Hashtags, Links

$$G(i) = \sum_{c \in \mathbb{C}} \mathbb{P}^2(i|c) = \sum_{c \in \mathbb{C}} \left(\frac{DF_i(c)}{DF_{\mathbb{T}}(i)} \right)^2 \qquad (1)$$

Extra-features

- Part-of-speech tagging, Named entities
- Tweet enrichment

Learning Strategies

- Specific training level, entity, domain, global
- Data sampling
- Monolingual and multilingual approaches

What about performances ?

Is tweet content sufficient ?

- Basic content-based classifiers using TFxIDFxGini over bag-of-N-grams
- No specific training, parameters tuning, or features selection

Dimensions classification performances

System	F-Score	Accuracy
Cosine	**.505**	.741
CRF	.492	**.771***
Best_System	*.473*	*.731*
SVM	.467	.733

Priority detection performances

System	F-Score	Accuracy	F(R&S)
Best_System	**.571**	.636	**.335**
SVM	.553	**.643**	.304
CRF	.554	.633	.318
Baseline	*.512*	*.570*	*.274*
Cosine	.566	.637	.260

RepLab: Author Profiling

RepLab Author Profiling issues

Profiling Twitter Authors

- Identify Opinion Makers
- Categorize profile according to their activity
- See PAN for gender and age identification

Data

- Large profile collection (7,000 tweets)
- 2 economic domains Automotive, Banking
- Profile comprises the last 600 posted tweets

Profile representation

- A profile is a bag of tweets where each one is tagged
  Tweets can be bag of words or set of features
- A profile is a bag of words or features (user as document)

Features

- Bag of N-grams representation (TF, IDF, purity index)

Strategies

- Majority vote over the profile
- Monolingual and domain specific model VS global model

You are what your network is

Features

- SNA features (followers, followees, etc.)
- Writing behaviour (retweets, hashtag, links)
- Public profile (description and personal data)
- External data (klout, kred)

Features issues

- Features names
- Features relevance to the problem

Twitter Profile

Julio's profile

- From these data, is Julio an influent Twitter user ?

What about performances ?

Is tweet content sufficient ?

- Basic content-based classifiers using bag-of-N-grams with TFxIDFxGini
- No specific training, parameters tuning, or features selection

Author ranking performances

System	Automotive	Banking	Avg MAP
Cosine	**.803**	.626	**.714**
LIA_new	.764	**.652**	.708
Best_System	.721	.410	.565
Baseline	.370	.385	.378
Klout	.304	.275	.289

- LIA_New is a tuned KNN*
- Best_System use hot topics information
- Followers baseline
- Raw features (Tweets, followees etc.) are under Klout

What about performances ?

Is tweet content sufficient ?

- Basic content-based classifiers using TFxIDFxGini over bag-of-N-grams
- No specific training, parameters tuning, or features selection
- Raw features used

Influence detection classification

System	Automotive	Banking	Avg F
Cosine	**.833**	*.751*	**.792**
LIA_new	.702	.726	.714
Best_System	.696	.693	.694
Baseline	.500	.500	.500

Wrap up !

What's been done in RepLab

- Machine Learning
- Tweet expansion
- Data sampling
- Active Learning

Some conclusions

- SVM-based methods appear to be sensitive to the training diversity
- We are still looking to the ultimate feature/approach
- Complex techniques are not significantly better
- Simple statistical NLP performs well
- Evaluation depends on the system ability to understand experts failures
- Evaluation is senstive to duplicate contents

Wrap up !

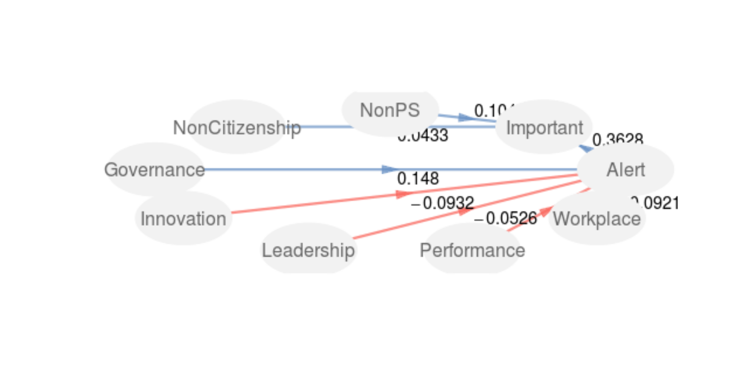
What's not been done in RepLab

- Reputation modeling
  Which aspects is the most important ?
  What is the impact of the company's innovation on its reputation ?
- Reputation-oriented summaries
  What is said on a particular topic ?

What's next ?

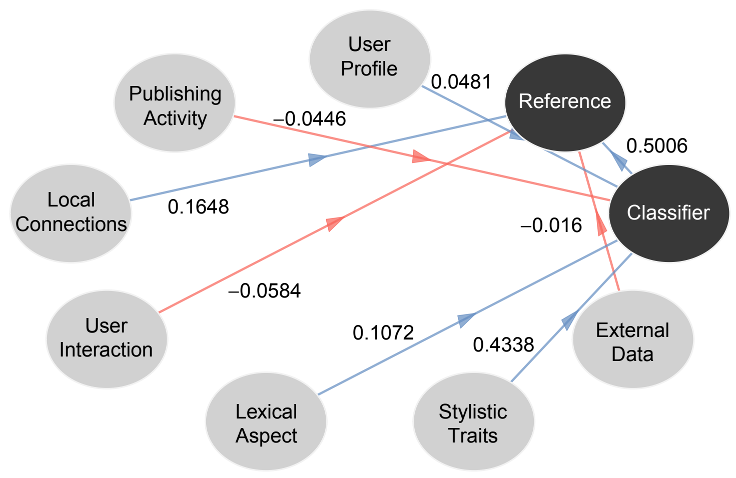
- Profile summarization in progress
- Active Learning with experts

Music Reputation Modeling with reference Alert



- Missing domain specific training material
- Dimensions inferred from Automotive and Banking
- Almost no correlation !

Features behaviour regarding influence for banking



Thank you !

Contact:

- jean-valere.cossu@univ-avignon.fr
- www.jeanvalerecossu.fr

References

- **Detecting Real-World Influence Through Twitter**
  Cossu J-V., Dugue N. and Labatut V. @ENIC 2015
- **Automatic Classification and PLS-PM Modeling for Profiling Reputation of Corporate Entities on Twitter**
  Cossu J-V., SanJuan E., Torres J-M. and El-Bèze M. @NLDB 2015
- **Tweet Expansion Method for Filtering Task in Twitter.**
  Karisani P., Oroumchian F. and Rahgozar M. @CLEF 2015
- **Overview of the 3rd Author Profiling task at PAN 2015.**
  Rangel F., Rosso P., Potthast M., Stein B., and Daelemans W., D. @CLEF 2015