



Emotional Social Signals for Search Ranking

Ismail Badache, Mohand Boughanem

► To cite this version:

Ismail Badache, Mohand Boughanem. Emotional Social Signals for Search Ranking. 40th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2017), Aug 2017, Tokyo, Japan. , pp.1053-1056. <hal-03154389>

HAL Id: hal-03154389

<https://hal.science/hal-03154389v1>

Submitted on 28 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



Emotional Social Signals for Search Ranking

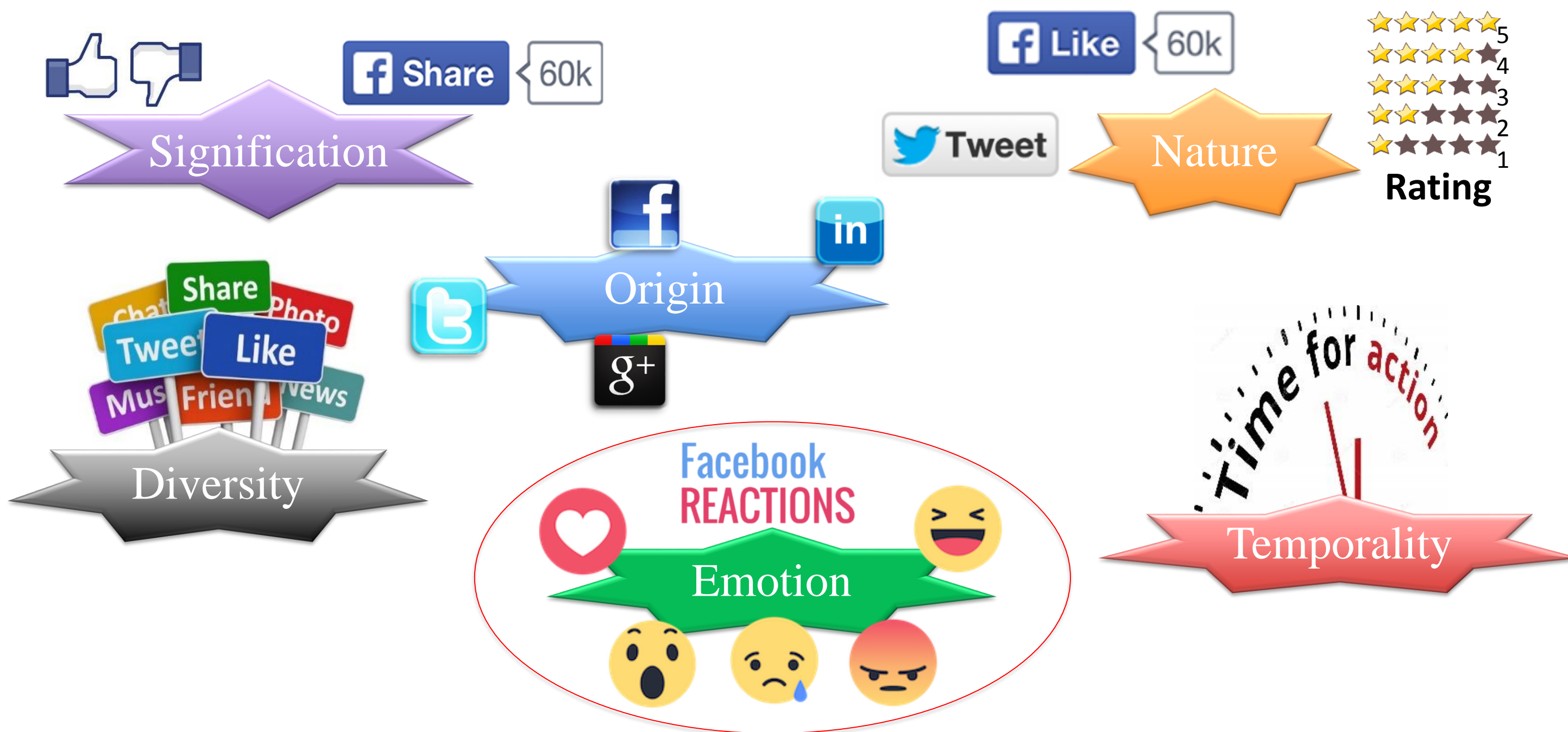


Ismail Badache and Mohand Boughanem
IRIT - Paul Sabatier University, Toulouse, France
{Badache, Boughanem}@irit.fr

1. Introduction

► General Context:

- Exploiting social signals to enhance a search.



► Research Questions:

- How users use these reactions (emotional signals) to interact with resources?
- What is the impact of Facebook reactions on IR process?

2. Impact of Emotional Signals

- Using any textual model to estimate the relevance of document D to a query Q .

$$RSV(D, Q) =_{Rank} P(D) \cdot RSV_{textual}(Q, D) \\ =_{Rank} P(D) \cdot \prod_{w_i \in Q} RSV_{textual}(w_i | D) \quad (1)$$

$P(D)$ is a document prior. w_i represents words of query Q .

- The priors are estimated by a counting of reactions r_i associated with D .

$$P(D) = \prod_{r_i \in R} P(r_i) \quad (2)$$

- Smoothing $P(r_i)$ by collection C using *Dirichlet* :

$$P(D) = \prod_{r_i \in R} \left(\frac{Count(r_i, D) + \mu \cdot P(r_i | C)}{Count(r_\bullet, D) + \mu} \right) \quad (3)$$

Where:

- $P(D)$ represents the a priori probability of D .
- $Count(r_i, D)$ represents number of occurrence of reaction r_i on resource D .
- r_\bullet is the total number of reactions.

3. Previous Work

► Previous contributions:

- Evaluating the impact of several *signals (individually & grouped)*.
[Badache and Boughanem, IIR'14]
- Evaluating the impact of signals *freshness* and *resource age*.
[Badache and Boughanem, ECIR'15]
- Evaluating the impact of signals *diversity* and *signals quality*.
[Badache and Boughanem, SIGIR'15]
- Conducting an *extensive experiments* on *SBS* and *IMDb* datasets.
[Badache and Boughanem, CHIIR'17]
- [Badache *et al.*, CLEF'15]

► Datasets:

- INEX Social Book Search 2015, 208 topics and their qrels.
- INEX IMDb 2011, 30 topics and their qrels.

► References:

- IIR'14 «Social Priors to Estimate Relevance of a Resource».
- ECIR'15 «Document Priors Based on Time-Sensitive Social Signals».
- CLEF'15 «Multimodal Social Book Search».
- SIGIR'15 «A Priori Relevance Based on Quality and Diversity of Social Signals».
- CHIIR'17 «Fresh and Diverse Social Signals: Any Impacts on Search?».

Table 1. Results of P@{10, 20}, nDCG and MAP (IMDb)

Models	P@10	P@20	nDCG	MAP
Baseline (A) : Without Priors				
ML	0.3700	0.3403	0.4325	0.2402
Baseline (B) : Without Considering Time				
Like	0.3938	0.3620	0.5130	0.2832
Share	0.4061	0.3649	0.5262	0.2905
Comment	0.3857	0.3551	0.5121	0.2813
TotalFacebook	0.4209	0.4102	0.5681	0.3125
Tweet	0.3879	0.3512	0.4769	0.2735
+1	0.3826	0.3468	0.5017	0.2704
Bookmark	0.3730	0.3414	0.4621	0.2600
Share (LIn)	0.3739	0.3432	0.4566	0.2515
All Criteria	0.4408	0.4262	0.5974	0.3300
(b) With Considering Publication Date T_D				
Like ^{T_D}	0.4091*	0.3620*	0.5308*	0.2907*
Share ^{T_D}	0.4177*	0.3721*	0.5544*	0.2989*
Comment ^{T_D}	0.3912*	0.3683*	0.5285*	0.2874*
TotalFacebook ^{T_D}	0.4302	0.4258	0.5827	0.3200
Tweet ^{T_D}	0.3918*	0.3579*	0.4903*	0.2779*
+1 ^{T_D}	0.3900	0.3511	0.5246	0.2748
Bookmark ^{T_D}	0.3732	0.3427	0.4671	0.2618
Share ^{T_D} (LIn)	0.3762	0.3449	0.4606	0.2542
All Criteria ^{T_D}	0.4484*	0.4305*	0.6200*	0.3366*
With Considering Diversity Div				
TotalFacebook ^{Div}	0.4227*	0.4187*	0.5713*	0.3167*
All Criteria ^{Div}	0.4463*	0.4318*	0.6174*	0.3325*

Table 2. Results of P@{10, 20}, nDCG and MAP (SBS)

Models	P@10	P@20	nDCG	MAP
Baseline (A) : Without Priors				
ML	0.0607	0.0559	0.1620	0.0527
Baseline (B) : Without Considering Time				
Like	0.0857	0.0689	0.1864	0.0741
Share	0.0901	0.0711	0.1900	0.0872
Comment	0.0799	0.0678	0.1807	0.0701
TotalFacebook	0.0958	0.0810	0.1937	0.0892
BA_Rating	0.0730	0.0559	0.1748	0.0620
Log_Tag	0.0770	0.0531	0.1742	0.0610
All Criteria	0.0973	0.0787	0.1981	0.0900
(a) With Considering Action Date T_a				
BA_Rating ^{T_a}	0.0941**	0.0732**	0.1904**	0.0885**
(b) With Considering Publication Date T_D				
Like ^{T_D}	0.0891*	0.0708*	0.1900*	0.0873*
Share ^{T_D}	0.0917*	0.0796*	0.1947*	0.0903*
Comment ^{T_D}	0.0881*	0.0711*	0.1882*	0.0777*
TotalFacebook ^{T_D}	0.0957**	0.0873**	0.1959**	0.0928**
BA_Rating ^{T_D}	0.0790*	0.0695*	0.1808*	0.0685*
Log_Tag ^{T_D}	0.0782*	0.0599*	0.1771*	0.0666*
All Criteria ^{T_D}	0.1078**	0.0973**	0.2080**	0.0986**
With Considering Diversity Div				
TotalFacebook ^{Div}	0.0960*	0.0840*	0.1945*	0.0907*
All Criteria ^{Div}	0.1011*	0.0915*	0.2031*	0.0952*

4. Experimental Evaluation

► Data:

- Collected from Facebook between March - June, 2016.

Table 3. Reactions audiences in some IMDb Box Office 2016

Movie	Love	Haha	Wow	Angry	Sad
<i>Captain America</i>	91%	9%	0%	0%	0%
<i>The Jungle Book</i>	88%	8%	2%	0%	2%
<i>Conjuring 2</i>	87%	1%	7%	2%	3%
<i>Warcraft</i>	81%	3%	14%	2%	0%

Table 4. Reactions audiences in some International Media

Media	Love	Haha	Wow	Angry	Sad
<i>Washington Post</i>	63%	12%	11%	8%	6%
<i>New York Time</i>	61%	13%	12%	8%	6%
<i>Fox News</i>	60%	14%	10%	8%	8%
<i>Guardian</i>	52%	9%	3%	25%	11%

Fox News

Sydney Morning Herald

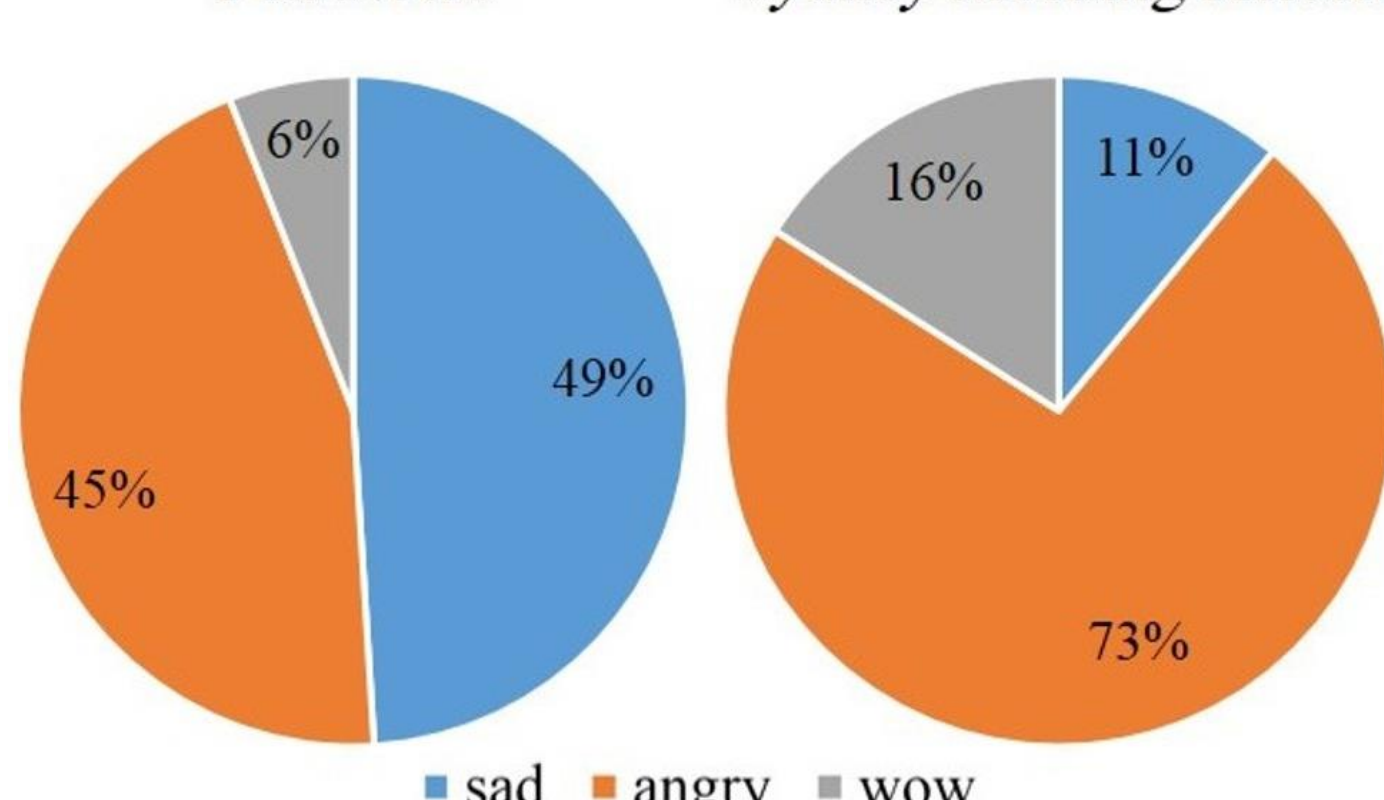


Figure 1. Some reactions to March 22 Brussels Attack

Table 5. Some emotionally-aware topics

id	Topics
Positively-aware topics	
2011014	romance movies by Richard Gere or George Clooney
2011102	best movies award "James Cameron"
2011116	action biker
2011121	spoofing comedies
Negatively-aware topics	
2011122	chernobyl
2011123	worst actor century
2011134	movies of actors death during filming
2011120	Vietnam war true story

	2011014	2011102	2011116	2011121
Love	0,9112	0,9401	0,5812	0,6629
Wow	0,1828	0,812	0,7306	0,3712
Haha	0,071	0,0144	0,3271	0,8602
Angry	0,1072	0,07	0,0723	0,0621
Sad	0,1013	0,231	0,0611	0,0179
	2011122	2011123	2011134	2011120
Love	0,0512	0,0111	0,0558	0,1137
Wow	0,124	0,0773	0,2174	0,1708
Haha	0,0114	0,1705	0,1302	0,0212
Angry	0,761	0,7126	0,7553	0,6725
Sad	0,68	0,6021	0,8947	0,8331

Table 6. P@10 related to some emotionally-aware topics

► Dataset:

- INEX IMDb 2011, 30 topics and their qrels.

Table 7. Results of P@{10, 20}, nDCG and MAP

IR Models	P@10	P@20	nDCG	MAP
Baseline: Textual Model				
LM	0.3700	0.3403	0.4325	0.2402
Baseline: Facebook Like				
Like	0.3938	0.3620	0.5130	0.2832
Facebook Reactions Individually				
Love	0.4122*	0.3702*	0.5300*	0.2978*
Haha	0.3900	0.3624	0.5100	0.2766
Wow	0.4031*	0.3755*	0.5203*	0.2889*
Sad	0.3800	0.3505	0.4811	0.2700
Angry	0.3111	0.2814	0.3421	0.1601
Facebook Reactions Combinations				
(Love, Haha) ⁺	0.4187*	0.3801*	0.5555*	0.2991*
(Sad, Angry) ⁻	0.3021	0.2614	0.3167	0.1574
(Love, Haha, Wow) ⁺	0.4275*	0.4112*	0.5773*	0.3168*

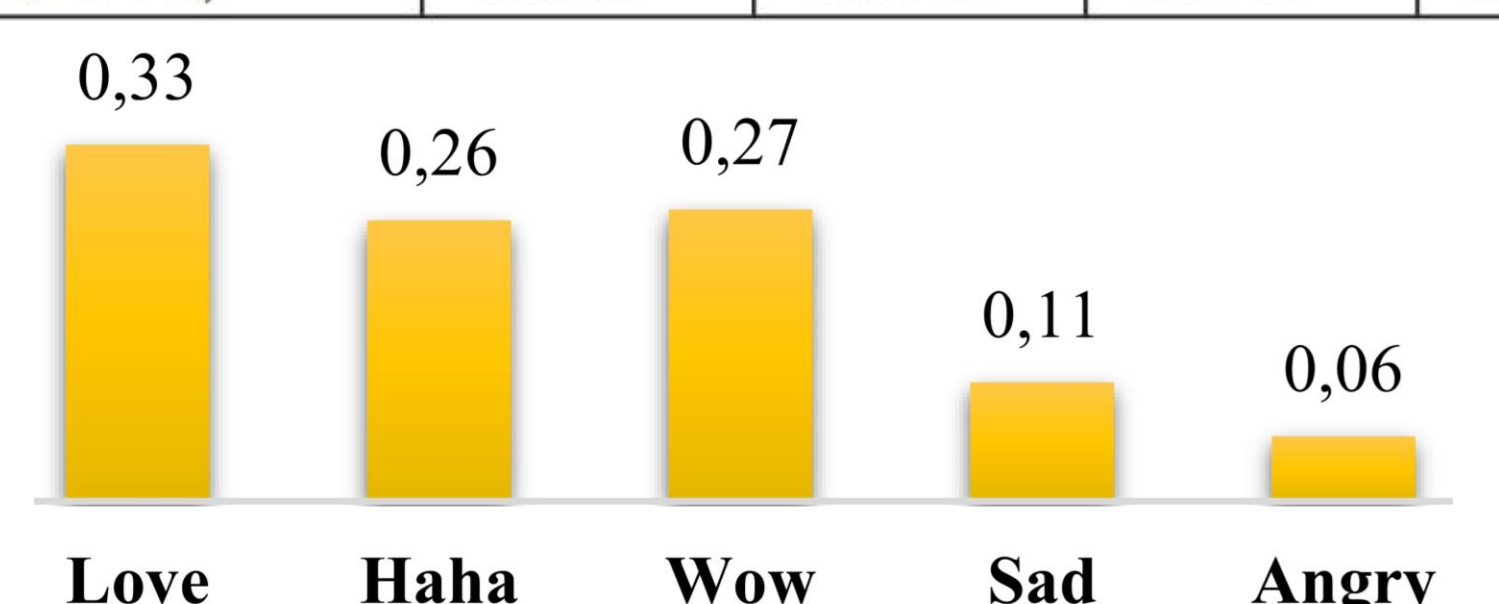


Figure 2. Rho correlation between relevance and reactions