Automatic Identification of Eyewitness Messages on Twitter during Disasters

Kiran Zahra

University of Zurich, Switzerland

New York plane crash: Twitter breaks the news, again - The Daily Telegraph
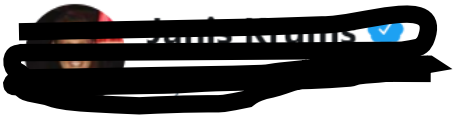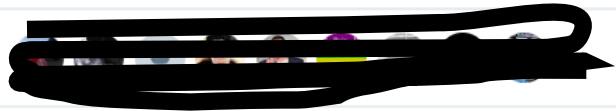

http://twitpic.com/135xa - There's a plane in the Hudson. I'm on the ferry going to pick up the people. Crazy.

12:36 PM - 15 Jan 2009

266 Retweets **944** Likes

💬 29   ⟲ 266   ♡ 944   ✉

2 women save a dog during a flood in India

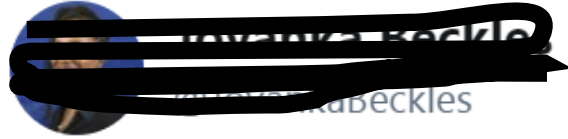**2 women save a dog during a flood in India • r/HumansBein...**

20980 points and 443 comments so far on reddit

reddit.com

10:25 PM - 17 Aug 2017

💬   🔁   ♡   ✉

...Beckles

In 2010, the Citizens United decision was an earthquake in American politics. For the first time, corporations and billionaire PACs could spend unlimited amounts of money to elect politicians that serve their interests, not yours.

Eyewitness Types

- Direct eyewitness
- Indirect eyewitness
- Vulnerable direct eyewitness

Direct Eyewitness : Earthquake

big midnight earthquake and aftershocks now
🙃

10:04 AM - 1 Aug 2017

11 Likes
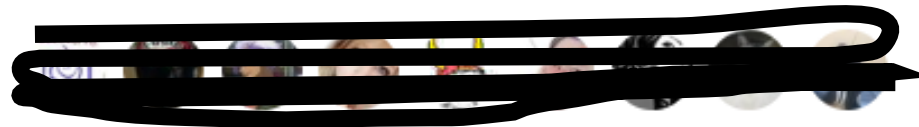
💬 3   🔁   ♡ 11   ✉

Indirect Eyewitness : Earthquake

F*cking hell... my wife and kids are in Tokyo and they're in the middle of an earthquake Jesus Murphy just how crap can one day get?

10:11 AM - 1 Aug 2017

Vulnerable Direct Eyewitness : Flood

Why am I always napping when a flash flood warning comes on to my phone? #scared

10:59 AM - 2 Aug 2017

4 Likes

💬 2   🔁   ♡ 4   ✉

Domain expert features - Direct eyewitness

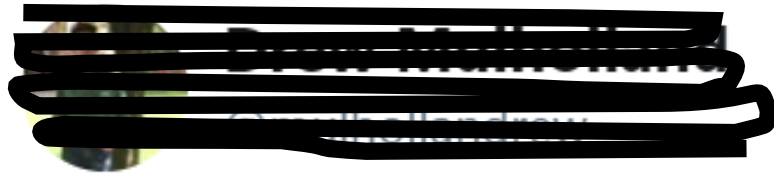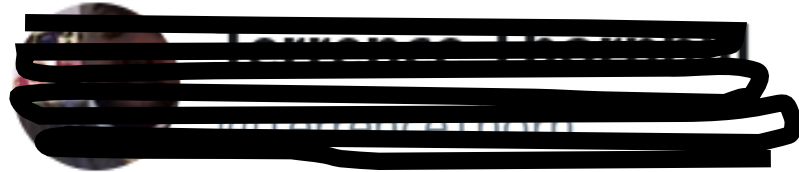| No.	Characteristic	Examples	EQ	FLD	HUR
(1)	Reporting small details of vicinity	window shaking, water in basement	X	X	X
(2)	Words indicating perceptual senses	seeing, hearing, feeling	X	X	X
(3)	Reporting impact of disaster	raining, school canceled, flight delayed	X	X	X
(4)	Words indicating intensity of disaster	intense, strong, dangerous, big	X	X	X
(5)	First person pronouns and adjectives	I, we, me	X	X	X
(6)	Personalized location markers	my office, our area	X	X	X
(7)	Exclamation and question marks	! , ?	X	X	X
(8)	Special/swear words	wtf, omg, s**t	X	X	X
(9)	Mention of a routine activity	sleeping, watching a movie	X	X	-
(10)	Time indicating words	now, at the moment, just	X	-	-
(11)	Short tweet length	one or two words	X	-	-
(12)	Caution and advices for others	watch out, be careful	-	X	X
(13)	Mention of disaster locations	area and street name, directions	-	X	X

Domain expert features - Indirect eyewitness

No.	Characteristic	Examples	EQ	FLD	HUR
(1)	Mention of locations or people the author knows	mom, dad, hometown	X	X	X
(2)	First person adjective	my, our	X	X	X
(3)	Expressing emotions	thoughts, worry, relief	X	X	X
(4)	Reporting safety, damage, missing	missing, safe	X	X	X

Domain expert features - Vulnerable direct eyewitness (similar)

No.	Characteristic	Examples	EQ	FLD	HUR
(1)	Reporting small details of vicinity	window shaking, water in basement	X	X	X
(2)	Words indicating perceptual senses	seeing, hearing, feeling	X	X	X
(3)	Reporting impact of disaster	raining, school canceled, flight delayed	X	X	X
(4)	Words indicating intensity of disaster	intense, strong, dangerous, big	X	X	X
(5)	First person pronouns and adjectives	I, we, me	X	X	X
(6)	Personalized location markers	my office, our area	X	X	X
(7)	Exclamation and question marks	! , ?	X	X	X
(8)	Special/swear words	wtf, omg, s**t	X	X	X
(9)	Mention of a routine activity	sleeping, watching a movie	X	X	-
(10)	Time indicating words	now, at the moment, just	X	-	-
(11)	Short tweet length	one or two words	X	-	-
(12)	Caution and advices for others	watch out, be careful	-	X	X
(13)	Mention of disaster locations	area and street name, directions	-	X	X

Domain expert features - Vulnerable direct eyewitness (distinct)

No.	Characteristic	Examples	EQ	FLD	HUR
(1)	Warnings and alerts about expected disasters	flash flood warnings	-	X	X
(2)	Associating warnings with current weather situation	flash flood alert with rain	-	X	X
(3)	Expressing emotions	hate, disgust, anger, scare	-	X	X

Machine learning classification

- Eyewitness
- Non-Eyewitness
- Noise

Classification models

- Textual features (BOW) – Baseline
- Domain-expert features
- Text and domain-expert features
- Text and domain-expert features with class balancing

Earthquake classification results

Text-based features (baseline)

Category	Precision	Recall	F-score	Class Dist.
Eyewitness	0.878	0.977	0.925	1600
Non-eyewitness	0.893	0.585	0.707	200
Don't know	0.629	0.280	0.388	200

Domain-expert features

Category	Precision	Recall	F-score	Class Dist.
Eyewitness	0.871	0.969	0.917	1600
Non-eyewitness	0.787	0.645	0.709	200
Don't know	0.333	0.095	0.148	200

Domain-expert + text features

Category	Precision	Recall	F-score	Class Dist.
Eyewitness	0.865	0.987	0.922	1600
Non-eyewitness	0.912	0.620	0.738	200
Don't know	0.641	0.125	0.209	200

Domain-expert + text features with class balancing

Category	Precision	Recall	F-score	Class Dist.
Eyewitness	0.892	0.966	0.927	1600
Non-eyewitness	0.932	0.793	0.857	400 (+100%)
Don't know	0.801	0.653	0.719	400 (+100%)

Wildfire classification results

Text-based features (baseline)			
Category	**Precision**	**Recall**	**F-score**	**Class Dist.**
Eyewitness	0.649	0.265	0.376	189
Non-eyewitness	0.857	0.941	0.897	1379
Don't know	0.748	0.708	0.728	432

Domain-expert features			
Eyewitness	0.703	0.339	0.457	189
Non-eyewitness	0.863	0.943	0.901	1379
Don't know	0.737	0.688	0.711	432

Domain-expert + text features			
Eyewitness	0.794	0.265	0.397	189
Non-eyewitness	0.867	0.946	0.905	1379
Don't know	0.730	0.731	0.731	432

Domain-expert + text features with class balancing			
Eyewitness	0.897	0.714	0.795	378 (+100%)
Non-eyewitness	0.753	0.727	0.740	432
Don't know	0.876	0.935	0.905	1379

Full Paper

- Zahra, Kiran, Muhammad Imran, and Frank O. Ostermann. "Automatic identification of eyewitness messages on twitter during disasters." *Information processing & management* 57.1 (2020): 102107.

Conclusion

- Finding firsthand and credible information during disasters is an important task for relief organizations

- Automatic methods to extract credible information from social media can lay the foundations of future communication