Teargas, Water Cannons and Twitter: A case study on detecting protest repression events in Turkey 2013

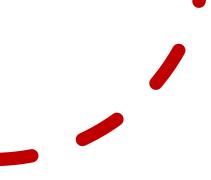
Fatma Elsafoury





Protest Repression

- "Any action by another group which raises the contender's cost of collective action" [Til78].
- We are interested in observable, coercive actions carried out by state agents against protesters.
 - Observable: seen by the public.
 - Coercive: include physical violence.
 - State agents: police.



Motivation & Research Proposal The available datasets have problems with:
Collected from News articles: coverage and accuracy bias [DB02, ESM03].
Hand labelling the data is time and labour consuming.

- Research proposal:
 - Collect events from social media as an alternative to News articles.
 - Automatic detection of protest repression events using ML to save time and labour.

Contribution

- Investigating Twitter as a data source for detecting protest repression (within the scope of our case study).
- Investigating crowdsourcing as a fast and cheap way to build a training dataset for the ML model (within the scope of our case study).

Case Study

Gezi Park Protest 2013 From 31/05/2013 to 30/06/2013.



Case Study

- The tweets were collected by SMPP using hashtags related to the protest.
- They collected 1,290,451 tweets in English.
- To use this dataset in training the ML model, we hire crowd workers to label the tweets.

Crowdsourcing Design

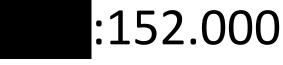
- Figure-Eight
- Hired 3 workers.
- Task limited to only workers from Turkey and with medium-level of experience.
- 116 test questions to eliminate spammers.
- 6693 tweets were labelled.

Questions

- 1. Is this tweet related to the Turkish Gezi park protest 2013?
- 2. Does this tweet report/discuss violent incident?

Crowdsourcing Results





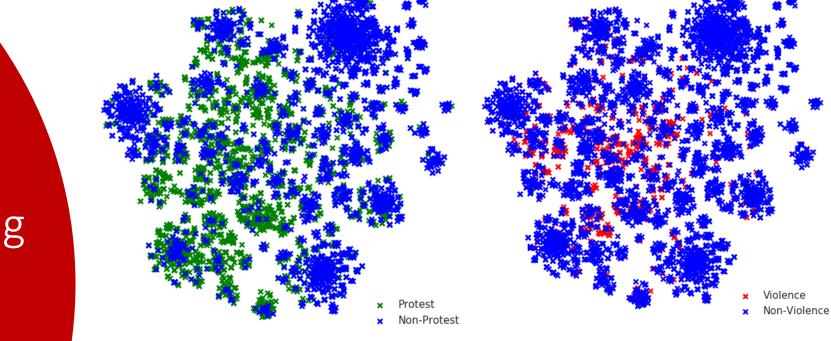
new owers, 30 new trees were planted to #GeziPark #Taksim after

it was cleaned of the #protestors"

crowdsourced tweets (labeled by Protest)

crowdsourced tweets (labeled by Violence)

Violence



Dataset	Size	Positive	Negative	
Protest	3860	39%	61%	
Violence	5247	6%	94%	

Crowdsourcing Full-agreed Results

ML experiment Design

- Preprocessing:
 - Remove user mentions, http links, hashtags, duplicated tweets.
 - Protest DS : 3666 tweets / 49% positive / 51% negative.
 - Violence DS : 4975 tweets / 6% positive / 94% negative.

• Feature:

- Word count, TF-IDF, W2V.
- Machine Learning Models:
 - SVM and MNB.

ML experiment Results

• Protest Classification:

	Feature	linear SVM $(C = 1)$	$MNB \ (\alpha = 2)$
	WC	0.895	0.856
	TF-IDF	0.896	0.846
•	WE	0.872	-

• Violence Classification:

Feature	linear SVM $(C = 10)$	MNB ($\alpha = 0.1$)
Count	0.8018	0.816
TF-IDF	0.8127	0.8189
WE	0.7689	-

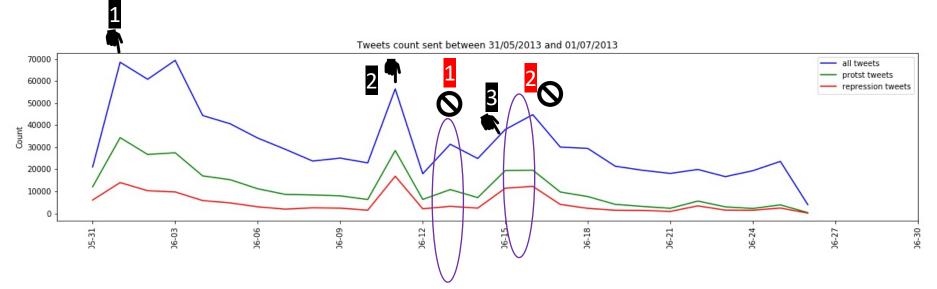
Results

- Models applied to the remaining 1,283,758 unlabelled tweets in our dataset.
- 67% of the protest related tweets don't report violence but 33% do report violence (protest repression incidents)

Model	Protest classification		Model	Violence classi	fication
	Positive	negative		Positive	negative
SVM + TF-IDF	36%	64%	MNB + TF-IDF	15%	85%

Data Analysis Tweets Timeline

Date	;	Event
31/0	5/2013	The beginning of the protest and the use of
		The beginning of the protest and the use of force by police including tear gas and water
		cannons against protesters.
11/0	6/2013	Police forces make an attempt to clear Gezi
		square by force.
15/0	6/2013	The square is successfully cleared from
		protesters.



so lear that water not clash non blenc the sear live tor budal thory in the section of the bir budal the section of the sectio 2013-05-31 -1.12 2013-06-01 2013-06-02 -1.15 0 22 0 22 0 32 0 13 0.01 0.23 0.11 2013-06-03 -1.17 0 15 0.03 0.06 0.05 2013-06-04 -0.53 0.40 011 015 001 004 004 2013-06-05 -0.49 0 10 0 00 0 03 0 13 2013-06-06 2013-06-07 2013-06-08 -0.29 0.01 0.01 2013.06.09 0 03 0 03 0 06 0 01 0 00 0 01 2013-06-10 0.02 0.03 0.02 0.01 0.01 0.03 0.01 0.04 0.02 0.01 0.01 0.02 0.01 0.00 0.01 0.00 0 99 0 92 0 35 1 20 0 36 1 12 0 31 0 20 0 42 0 23 0 12 0 26 0 37 0 09 0 33 0 46 0 25 0 21 0 31 0 05 0 04 0 35 2013-06-11 0.01 0.00 0.03 2013-06-13 -0.35 0.28 0.17 0.07 0.12 0.13 0.05 0.12 0.04 0.08 0.10 0.07 0.05 0.06 0.05 0.03 0.0 0.02 0.03 0.03 0.05 0.05 0.04 0.01 0.02 0.04 0.02 0.02 0.05 0.02 0.04 0.01 0.02 0.04 2013-06-14 1.22 0.63 0.10 0.73 0.38 0.44 0.27 0.43 0.40 0.30 0.30 0.41 0.28 0.10 0.31 0.08 0.20 0.16 0.24 0.12 0.12 0.99 0.32 0.07 2013-06-15 2013-06-16 2013-06-17 2013-06-19 -0 18 2013-06-20 -0.15 2013-06-21 0.02 0.01 2013-06-22 -0.39 0.26 0.01 0.03 0.03 2013-06-23 -0.14 0.13 0 00 0 01 0 01 2013-06-24 -0.05 0.05 0.00 0.00 0.01 2013-06-26 -0.01 0.01

Gas

Tear

-25

-2.0

-15 0

-1.0

Attack

Water

Riot

Clash

Force

Brutal

Report

Injury

Cannon

Violence

Data Analysis Tweets Timeline

Limitations & Future work

Limitations

- Ground Truth.
- The small dataset.
- The subjectivity of repression/violence.

Future work

- BERT
- Combine image with text classification for better detection.

Thanks For Listening!

Questions??





References

- [Til78] Charles Tilly. Collective violence in European perspective. 1978.
- [Ear03] Jennifer Earl. Tanks, tear gas, and taxes: Toward a theory of movement repression. Sociological theory, 21(1):44{68, 2003.
- [OBBC13] Isabel Ortiz, Sara L Burke, Mohamed Berrada, and Hernan Cortes. World Protests 2006-2013.2013.
- [DB02] Christian Davenport and Patrick Ball. Views to a kill: Exploring the implications of source selection in the case of Guatemalan state terror, 1977-1995. Journal of conict resolution, 46(3):427{450, 2002}.
- ESM03] Jennifer Earl, Sarah A Soule, and John D McCarthy. Protest under re? Explaining the policing of protest. American sociological review, pages 581{606, 2003}.