The Role of Crowdsourcing in Fighting Online Misinformation

Gianluca Demartini

The University of Queensland, Australia

Presented at the
Conference for Truth and Trust Online 2020

Fighting Online Misinformation

- Check-whortiness
  - Deciding if some statement would benefit from fact-checking
- Fact-checking
  - A forensic process performed by expert journalists
- Truthfulness assessment/classification
  - Multi-class classification problem that a supervised ML model might be able to address
  - FEVER

Crowdsourcing Truthfulness Judgements

- ~600 MTurk US workers
- To assess truthfulness of
  - US political statements (Politifact)
  - non-US political statements (ABC)
- 3 scales (3, 6, and 100 levels)

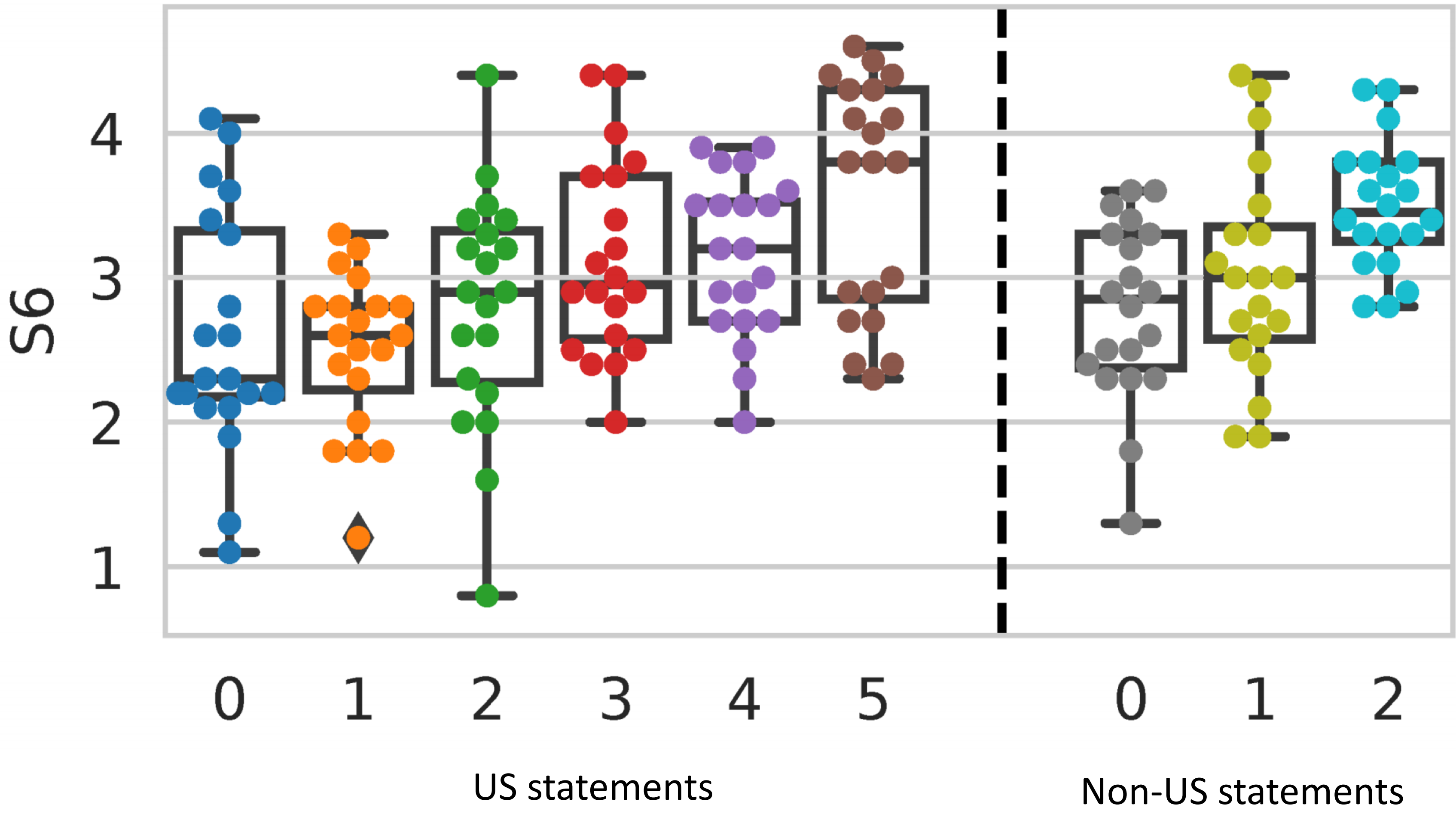
Table 1: Example of statements in the PolitiFact and ABC datasets.

	Statement	Speaker, Year
PolitiFact Label: mostly-true	"Florida ranks first in the nation for access to free prekindergarten."	Rick Scott, 2014
ABC Label: in-between	"Scrapping the carbon tax means every household will be $550 a year better off."	Tony Abbott, 2014

- All data:
- https://github.com/kevinRoitero/crowdsourcingTruthfulness

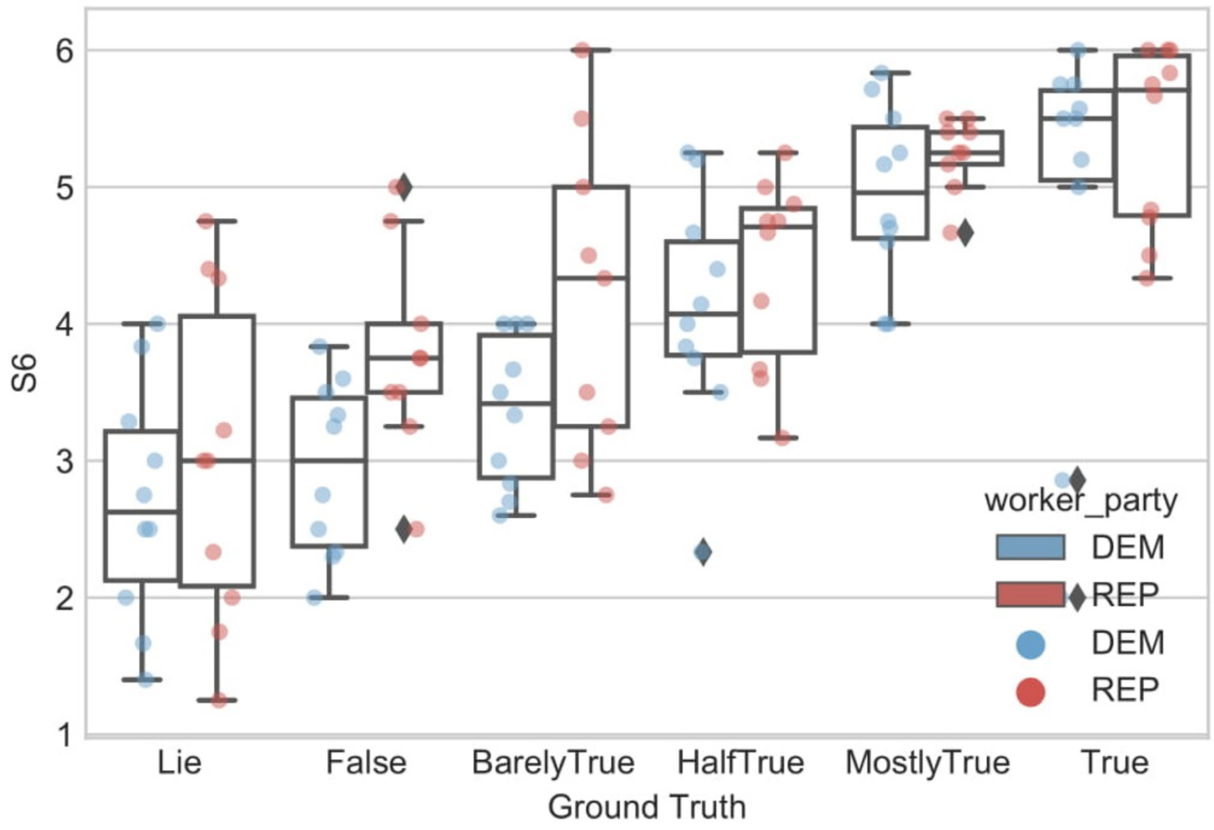
Kevin Roitero, Michael Soprano, Shaoyang Fan, Damiano Spina, Stefano Mizzaro and Gianluca Demartini. **Can The Crowd Identify Misinformation Objectively? The Effects of Judgments Scale and Assessor's Bias**. In: The 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2020)

Crowd Performance VS Expert Ground Truth



Political Bias

- Crowd workers who vote REP (red dots) are more likely to believe to statements by REP politicians

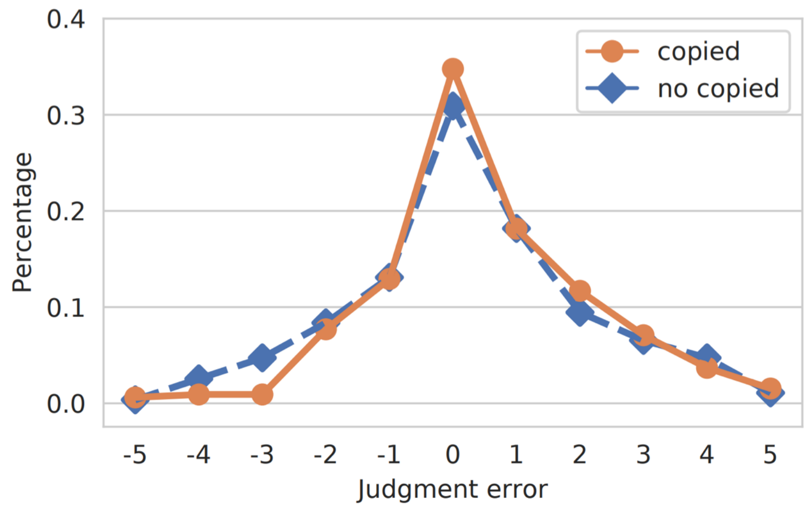


David La Barbera, Kevin Roitero, Damiano Spina, Stefano Mizzaro, and Gianluca Demartini. **Crowdsourcing Truthfulness: The Impact of Judgment Scale and Assessor Bias**. In: The 42nd European Conference on Information Retrieval (ECIR 2020). Lisbon, Portugal, April 2020.

Source of Support Evidence

URL	Percentage%
snopes.com	11.79%
msn.com	8.93%
factcheck.org	6.79%
wral.com	6.79%
usatoday.com	5.36%
statesman.com	4.64%
reuters.com	4.64%
cdc.gov	4.29%
mediabiasfactcheck.com	4.29%
businessinsider.com	3.93%

- We ask workers to
  - Search the web for supporting evidence
  - using a custom search engine where we remove Politifact pages from the results and
  - Provide a textual justification
- Workers who directly quote text from the selected web search result avoid underestimating the truthfulness of the statement



Kevin Roitero, Michael Soprano, Beatrice Portelli, Damiano Spina, Vincenzo Della Mea, Giuseppe Serra, Stefano Mizzaro, and Gianluca Demartini. **The COVID-19 Infodemic: Can the Crowd Judge Recent Misinformation Objectively?**. In: 29th ACM International Conference on Information and Knowledge Management (CIKM 2020)

Longitudinal COVID-19 Study

Table 1: Examples of COVID-19 fact-checked statements.

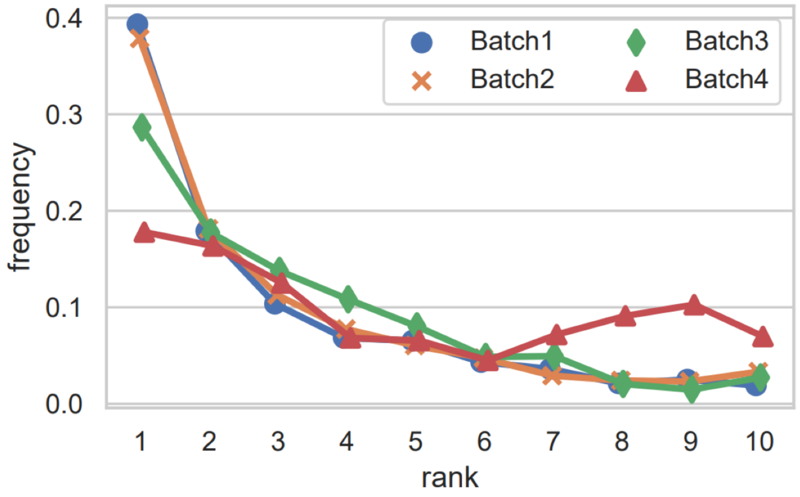
Statement	Source	Year	Label
"We inherited a broken test for COVID-19."	Donald Trump	2020	`pants-on-fire`
"Church services cannot resume until we are all vaccinated, says Bill Gates."	Bloggers	2020	`false`
"Says a 5G law passed while everyone was distracted with the coronavirus pandemic and lists 20 symptoms associated with 5G exposure."	Facebook Post	2020	`mostly-false`

Table 3: Experimental setting for the longitudinal study. All dates refer to 2020. Values reported are absolute numbers.

Date	Acronym	Number of Workers				Total
		Batch1	Batch2	Batch3	Batch4	
May	Batch1	100	–	–	–	100
June	Batch2	–	100	–	–	100
	$\texttt{Batch2}_{\texttt{from1}}$	29	–	–	–	29
July	Batch3	–	–	100	–	100
	$\texttt{Batch3}_{\texttt{from1}}$	22	–	–	–	22
	$\texttt{Batch3}_{\texttt{from2}}$	–	20	–	–	20
	$\texttt{Batch3}_{\texttt{from1or2}}$	22	20	–	–	42
August	Batch4	–	–	–	100	100
	$\texttt{Batch4}_{\texttt{from1}}$	27	–	–	–	27
	$\texttt{Batch4}_{\texttt{from2}}$	–	11	–	–	11
	$\texttt{Batch4}_{\texttt{from3}}$	–	–	33	–	33
	$\texttt{Batch4}_{\texttt{from1or2or3}}$	27	11	33	–	71
	$\texttt{Batch}_{\texttt{all}}$	100	100	100	100	400

Changes over time

- There is a significant difference in the quality of new workers from the different batches
  - Some statements (end of March and April) are the most difficult to assess
  - Time elapsed since the statement was made has no impact on crowd judgment quality
- Search results change over time and selected supporting URLs are found lower in the search engine result page



Hybrid Human-AI Approaches to Fighting Online Misinformation

- Crowd workers provide reliable (but not perfect) truthfulness labels
- AI can provide reliable (but not perfect) truthfulness labels
- Experts can provide perfect truthfulness labels and justifications

- Can we leverage them all to work effectively and at scale?

Gianluca Demartini, Stefano Mizzaro, and Damiano Spina. **Human-in-the-loop Artificial Intelligence for Fighting Online Misinformation: Challenges and Opportunities**. In: Data Engineering Bulletin, September 2020 issue.

Open Research Questions

- Who should do what?
  - Task allocation models
  - Cascade models: First AI to label at scale and quickly, then experts to “slowly” check the most important ones
- Urgency vs effectiveness
  - Identify difficult statements for expert to check and let “easy” ones for non-experts to label
- How would experts actually work when embedded in such a new framework
  - Trust in the hybrid system
  - Giving up levels of control: need for self-explainable human-in-the-loop AI tools

Conclusions

- There is a need to scale efforts to fight the growing issue of online misinformation
- Using AI and crowdsourcing can, in some cases, complement expert efforts
- A combined expert-AI-crowd approach could provide the best scale/quality/urgency trade-off