

A Dashboard for Mitigating the COVID-19 Misinfodemic

Zhengyuan Zhu, Kevin Meng, Josue Caraballo, Israa Jaradat, Xiao Shi, Zeyu Zhang, Farahnaz Akrami, Haojin Liao, Fatma Arslan, Damian Jimenez, Mohammed Samiul Saeef, Paras Pathak, Chengkai Li

Presenter: Zhengyuan Zhu

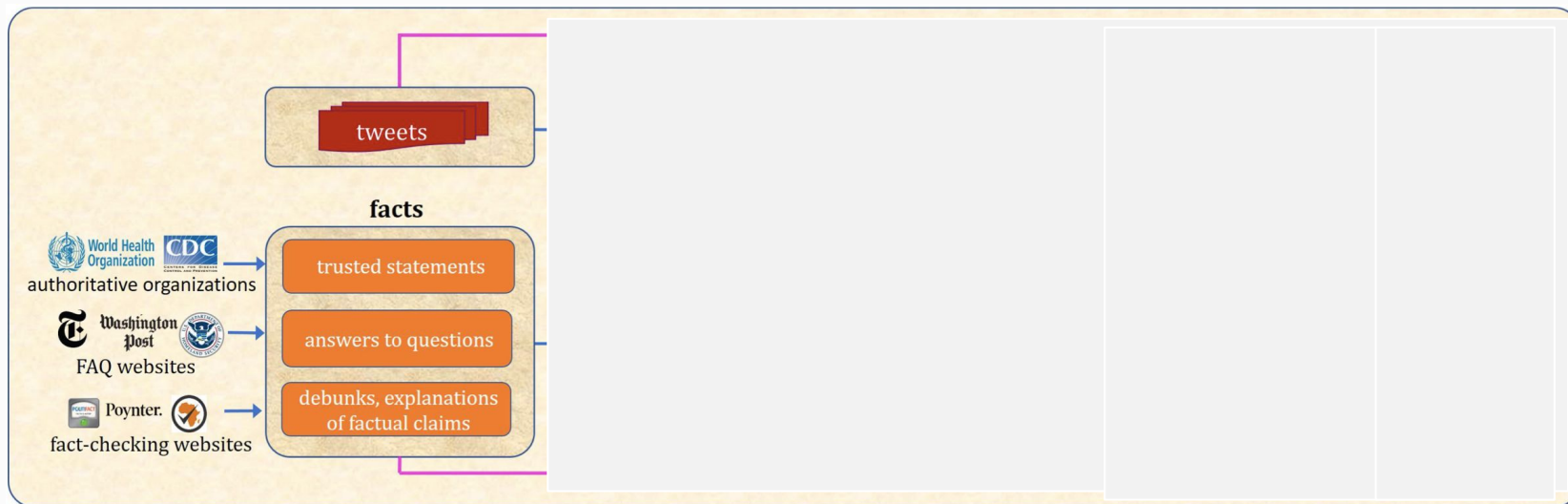
Outline

- **Project Goal**
- Project Demo
- Model Architecture
- Experiments
- Findings
- Conclusion

Project Goal

- **Build a dashboard displays the most prevalent facts and debunks of misinformation** among Twitter users in any user selected U.S. geographic region.
- **Facts and debunks:** Curate facts and debunks come from a catalog of COVID-19.
- **Misinformation:** Social media tweets hold a disagree stance towards a fact/debunk.
- **Example:**
 - **Tweet:** Coronavirus cannot be passed by dogs or cats but they can test positive.
 - **Fact:** There has been no evidence that pets such as dogs or cats can spread the coronavirus.
 - Stance: **agree**

Model structure



- Dataset:
 - Tweet: using a collection of approximately 250 million COVID-19 related tweets.
 - Facts: manually curated a catalog, which currently has 9,512 entries from 21 credible websites
- Methods: **two-stage pipeline for detecting misinformation**
 - Claim Matching: Sentence BERT
 - Stance Detection: Stance BERT

Experiments

- Stance BERT performance on FNC-1 Dataset:

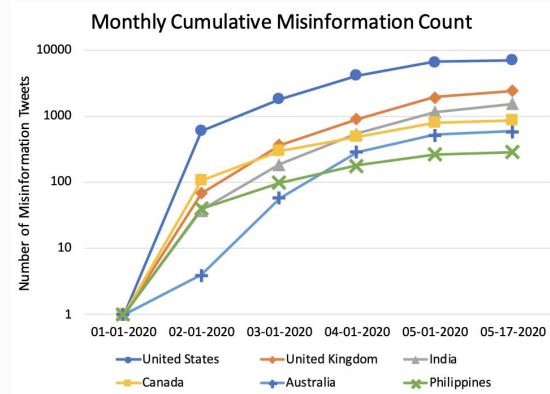
Model	F1 Score			
	agree	discuss	disagree	macro
Stance-BERT _{window}	0.65	0.45	0.84	0.65
Stance-BERT _{trunc}	0.66	0.41	0.82	0.63
(Xu et al., 2018)	0.55	0.15	0.73	0.48

- Stance BERT performance on COVID-19 Dataset:

Metric	@5	@10	@20	@50	@100
Precision	0.80	0.80	0.70	0.56	0.52
NDCG	0.62	0.72	0.78	0.81	0.83

Findings

- Top 6 countries with the most misinformation tweets:



- Correlation between the percentage of confirmed/deceased/recovered cases and the percentage 578 of misinformation tweets.

Country	Confirm	Death	Recover
United States	0.763	0.738	0.712
United Kingdom	0.862	0.833	-
India	0.794	0.798	0.755
Canada	0.706	0.667	0.663
Australia	0.954	0.922	0.887
Philippines	0.720	0.696	0.618

Conclusion

- In this demonstration, we introduce an information dashboard constructed in the context of our ongoing project regarding the COVID-19 misinfodemic.
- Going forward, we will focus on developing the dashboard at scale, including more comprehensive tweet collection and catalog discovery and collection.
- We will collect more labeled data for improving and evaluating our Stance-BERT models.
- We will also introduce more functions into the dashboard that are aligned with our project goal of studying the surveillance of, impact of, and intervention on COVID-19 misinfodemic.