Crowdsourced fact-checking: how is BirdWatch doing?

-
-
-

ABSTRACT

test

PVLDB Reference Format:
–. BirdWatch. *PVLDB*, 1(1): 2508 - 2521, 2021.
DOI: https://doi.org/XX

1. INTRODUCTION

Several papers study crowd-sourced fact checking, but how effective is it "in the wild"? The BirdWatch initiative provides us data to evaluate the effectiveness of this effort over multiple dimensions. In this setting, any user can create a *check* for a tweet (providing some metadata about the annotations) and other users can up/down *vote* such check. Multiple user can check the same content independently [Paolo: to verify how twitter handle this cases, do they try to unify the votes?].

Dimensions over the checks, most of those apply also for analysis of the votes:

1. accuracy wrt ground truth (provided by us and/or professional fact checkers)
2. accuracy wrt computational methods
3. agreement among workers
4. dependency between assessment quality and evidence provided (e.g., use URLs, length/quality of text justifications)
5. effect of aggregation functions
6. effect of scales transformations (?)
7. (missing: workers background / bias)

Dimensions over the checkers:

1. check submitted
2. check voted
3. more ?

Immediate questions:

1. do the checkers debunk worth-checking tweets?
2. are the checks effective in debunking the tweets?
3. are the votes effective in championing the good checks?
4. can we characterize the checkers with a degree of trust?

Bigger questions:

1. assuming we can characterize the trust of the checkers (and of the fact checking methods), can we develop a cost model for these verification solution?
2. given full profile (trust/cost) for checkers and a budget, how to assign tasks to maximize the number of verified claims with a certain level of trust?

2. PROBLEM MODEL

3. METHODS

4. EXPERIMENTS

5. RELATED WORK

- Crowd quality, controlled environment, impact of bias/background (https://dl.acm.org/doi/10.1145/3397271.3401112)
- COVID claims, crowd quality, controlled environment, impact bias/background (https://dl.acm.org/doi/10.1145/3340531.3

6. CONCLUSION

7. REFERENCES

Proceedings of the VLDB Endowment, Vol. 1, No. 1
ISSN 2150-8097.
DOI: https://doi.org/XX