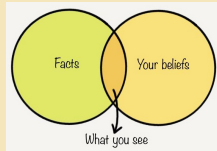# Accounting for Confirmation Bias in Crowdsourced Label Aggregation

Meric Altug Gemalmaz, Ming Yin

Department of Computer Science, Purdue University

Introduction

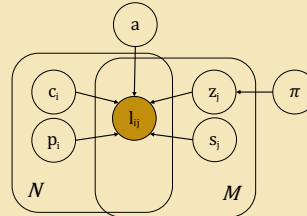**Crowd workers are prone to a wide range of biases!**

Confirmation Bias: Favoring information that confirms one's previously existing beliefs and values.

Label aggregation algorithms aim to increase crowd work quality by inferring ground truth from multiple annotations, but they seldom take worker bias into consideration…

Can we improve crowdsourcing label aggregation by explicitly account for worker's confirmation bias?

Our Approach

Use probabilistic graphic models to explicitly capture how worker's confirmation bias sneaks into annotations.
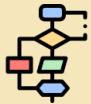
Design an expectation-maximization algorithm based on the proposed model to reduce the bias in aggregated labels.

Evaluation on both real-world datasets collected from Amazon Mechanical Turk and on synthetic datasets!

Our Model

Our Graphical Model:

Preferable Label Probability:

$$P(l_{ij} = 0 | c_i, p_i, s_j, z_j, a) = \frac{1}{e^{a[(1-p_i)(s_j-c_i)^2 + p_i z_j]}}$$

The closer c_i and s_j are to each other, the more likely worker **i** will provide the preferable label in task **j**.

Real-World Evaluation

Subjective Labeling Task: Label political statements related to gun control as either "Opinion" or "Factual".
- 12 statements (6 expressing liberal values and 6 expressing conservative values)
- 110 workers → 1320 labels

Existence of confirmation bias: Workers who self-report to have more liberal views are more likely to label statements expressing liberal values as factual.

Bias-aware label aggregation improves inference accuracy over baseline methods.
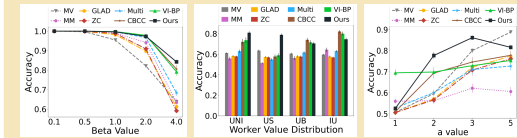
Simulation

The Impact of Confirmation Bias Degree
- $p_i \sim$ Beta$(1, \beta)$, $\beta \in \{0.1, 0.5, 1, 2, 4\}$
- The larger β is, the more workers are influenced by confirmation bias.

The Impact of the Distribution of Worker's Values
- $c_i \sim$ Beta(α, β)
- Uniform ($\alpha = \beta = 1$), U-shape($\alpha = \beta = 0.5$), Unbalanced($\alpha = 1$, $\beta = 2$), Inverse-U shape ($\alpha = \beta = 2$)
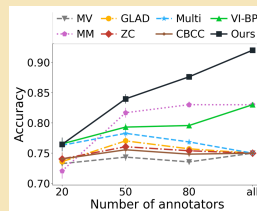
The Impact of Base Rate of the Preferable Label
- $a \in \{1, 2, 3, 5\}$
- The larger a is, the base chance for workers to provide the preferable label is lower.

Summary

- We propose a new label aggregation method that accounts for a worker's cognitive bias (confirmation bias).

- Both real-world data and simulation shows that our approach is advantageous, especially when the workers are biased and when the workers' values is more dispersed or even polarized.

- Accounting for cognitive bias in crowdsourcing label aggregation is feasible and promising!

Scan the code to check out our paper!

PURDUE UNIVERSITY

The 30th International Joint Conference on Artificial Intelligence, August 2021