# A Truthfulness Stance Map for 2024 Election-Related Factual Claims


UNIVERSITY OF TEXAS ARLINGTON

Caltech 2024 Election Integrity Project

Information and Misinformation in Elections: 2025 Conference

Zhengyuan Zhu  Zeyu Zhang  Haiqi Zhang  Supervisor: Chengkai Li

IDIR Innovative Data Intelligence Research Laboratory


## The Emphasis on Truthfulness Stance Detection

- We detect the stance taken by social media posts, specifically focusing on tweets, toward the truthfulness of election-related factual claims as the target.
- The truthfulness stances are classified into (1) **Positive stance**: The tweet believes the factual claim is true. (2) **Neutral stance or no stance**: The tweet expresses a neutral or no stance towards the factual claim's truthfulness. (3) **Negative stance**: The tweet believes the factual claim is false.

Tweet: This is beyond sick, it's called murder! 😡😡😡💔💔💔 — believe the claim is true → Claim: California introduces a new bill that would allow mothers to kill their babies up to 7 days after birth.

~~Tweet~~: The Assembly Bill 223, which is the bill in question, does not legalize infanticide. — believe the claim is false →

## Motivating Applications

- Use truthfulness stance as a tool to gauge public opinions expressed on social media toward election-related misinformation and comprehend how misinformation spreads and influences people's beliefs.
- Use truthfulness stance to identify demographics or communities that may be more susceptible to spreading misinformation.

## TSD-CT Dataset

**Claim collection:**

- Built an automated fact-check collection tool for seven websites.
- Collected 50,279 fact-checks, including factual claims from 2007 to 2023.

**Tweet collection:**

- Queried the Twitter API v2 using claim keywords to gather tweets related to PolitiFact's factual claims.
- Sanitized tweet-claim pairs. 5,793 pairs were prepared for annotation.

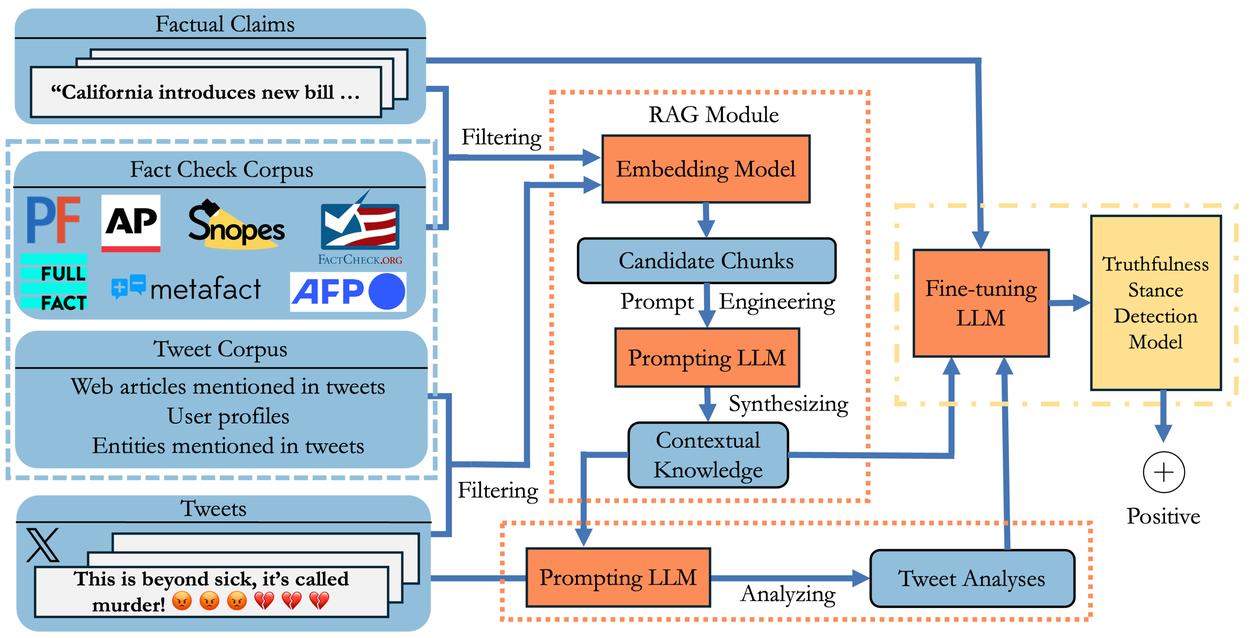
**Annotation collection:**

- Built a web annotation interface for collecting human annotations.
- This resulted in 3,105 completed pairs.

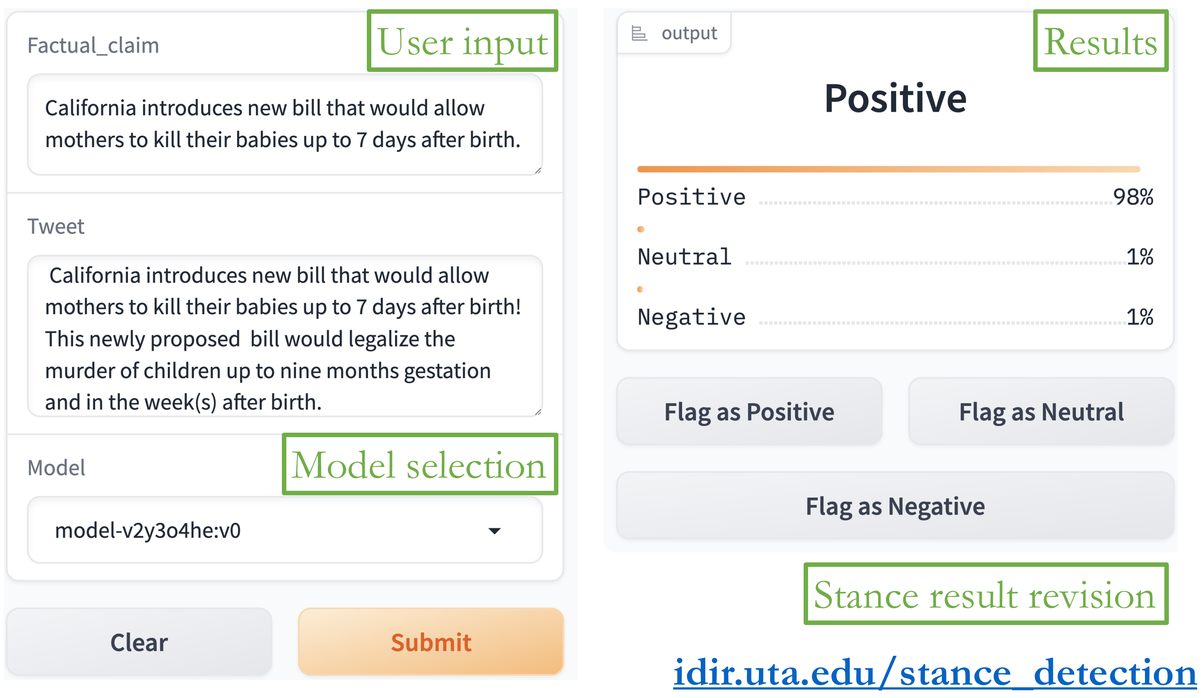
## Performance Comparison on TruthSD Dataset

| Model | $F_{\oplus}$ | $F_{\odot}$ | $F_{\ominus}$ | $F_M$ |
|---|---|---|---|---|
| BUT-FIT | 83.38 | 72.00 | 65.11 | 80.11 |
| BLCU_NLP | 85.37 | 71.43 | 63.29 | 73.36 |
| BERTSCORE+NLI | 88.68 | 72.53 | 81.04 | 80.75 |
| BART+NLI | 88.00 | 73.42 | 74.25 | 78.56 |
| TESTED | 84.09 | 72.37 | 67.90 | 74.75 |
| $RATSD_{Zephyr}$ | 88.67 | 77.38 | 80.28 | 82.10 |
| $RATSD_{GPT-3.5}$ | **93.27** | **80.24** | **87.90** | **87.13** |

## The RATSD Framework




## Truthfulness Stance Detection API




idir.uta.edu/stance_detection

## Case Study: Truthfulness Stance Map








## Acknowledgements

This work is partially supported by the National Science Foundation award # 2346261. The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC resources.