#### Weakly Supervised Learning for Analyzing Political Campaigns on Facebook

#### Tunazzina Islam, Shamik Roy, Dan Goldwasser

Department of Computer Science Purdue University, West Lafayette, IN 47907, USA

ICWSM 2023 *Date: June 5<sup>th</sup> – 8<sup>th</sup>, 2023* 



#### **Transformative** effect of Social Media

• Social media platforms: main channel for **political messaging**.



#### What is *Microtargeting*?

- Social media platforms: main channel for **political messaging**.
- Allows advertisers to **target** specific demographics and **adapt** their messaging based on their **reactions**.
  - Microtargeting: data-driven campaigning techniques exploiting the rich information collected by social networks about their users.



## Challenges

- Social media platforms: main channel for **political messaging**.
- Allows advertisers to **target** specific demographics and **adapt** their messaging based on their **reactions**.
  - Microtargeting: data-driven campaigning techniques exploiting the rich information collected by social networks about their users.
- Making this communication **transparent** is challenging.
  - Messaging is tightly coupled with its intended audience and often echoed by multiple stakeholders interested in advancing specific policies.



#### Goal

- Social media platforms: main channel for **political messaging**.
- Allows advertisers to **target** specific demographics and **adapt** their messaging based on their **reactions**.
  - Microtargeting: data-driven campaigning techniques exploiting the rich information collected by social networks about their users.
- Making this communication **transparent** is challenging.
  - Messaging is tightly coupled with its intended audience and often echoed by multiple stakeholders interested in advancing specific policies.
- Analyze and monitor the landscape of political advertising.
  - Our experiments: Analyze content supporting either the Biden-Harris or the Trump-Pence campaigns in U.S. 2020 presidential elections.





Targets males and age 25-34

Under the **Trump administration**, projections are for 20-30 percent *unemployment*, which would exceed the unemployment levels of the Great Depression. **Joe Biden** will *support the American economy by investing* \$400 *billion in products made by American workers*.

# Ť



Targets males and age 25-34

Under the **Trump administration**, projections are for 20-30 percent *unemployment*, which would exceed the unemployment levels of the Great Depression. **Joe Biden** will *support the American economy by investing* \$400 *billion in products made by American workers*.



Message Analysis Stance: **Pro-Biden** and **Anti-Trump** Issue: **Economy & taxes** 



Under the **Trump administration**, projections are for 20-30 percent *unemployment*, which would exceed the unemployment levels of the Great Depression. **Joe Biden** will *support the American economy by investing \$400 billion in products made by American workers*.



Message Analysis Stance: **Pro-Biden** and **Anti-Trump** Issue: **Economy & taxes** 



**Trump** is making *changes in medicare* that most seniors don't even know. These *changes are a good thing* and seniors are encouraged to check new trump rates.





Under the **Trump administration**, projections are for 20-30 percent *unemployment*, which would exceed the unemployment levels of the Great Depression. **Joe Biden** will *support the American economy by investing \$400 billion in products made by American workers*.



Message Analysis Stance: **Pro-Biden** and **Anti-Trump** Issue: **Economy & taxes** 



**Trump** is making *changes in medicare* that most seniors don't even know. These *changes are a good thing* and seniors are encouraged to check new trump rates.



Message Analysis Stance: **Pro-Trump** Issue: **Healthcare** 



Under the **Trump administration**, projections are for 20-30 percent *unemployment*, which would exceed the unemployment levels of the Great Depression. **Joe Biden** will *support the American economy by investing \$400 billion in products made by American workers*.



Message Analysis Stance: **Pro-Biden** and **Anti-Trump** Issue: **Economy & taxes** 



**Trump** is making *changes in medicare* that most seniors don't even know. These *changes are a good thing* and seniors are encouraged to check new trump rates.



Message Analysis Stance: **Pro-Trump** Issue: **Healthcare** 

#### **Our Work:**

1. Characterize the different stakeholders and analyze their content.

2. Analyze political messaging across different demographics.

10



	Dataset Details
	Weak Label Generation
• • •	Brief Description of the Framework
	Analysis of Messaging



	Dataset Details
	Weak Label Generation
• • • •	Brief Description of the Framework
	Analysis of Messaging

Facebook Ad Library API

#### Facebook Ad Library API

0.8 million political ads (English) focusing on U.S. 2020 presidential elections from January 2020 - October 2020.

- Facebook Ad Library API
- 0.8 million political ads (English) focusing on U.S. 2020 presidential elections from January 2020 October 2020.
- 35K ads with different contents, 5K unique funding entities, among them 537 explicitly mention candidate names and/or party affiliations, e.g., BIDEN FOR PRESIDENT, DONALD J. TRUMP FOR PRESIDENT, INC.

- Facebook Ad Library API
- 0.8 million political ads (English) focusing on U.S. 2020 presidential elections from January 2020 October 2020.
- 35K ads with different contents, 5K unique funding entities, among them 537 explicitly mention candidate names and/or party affiliations, e.g., BIDEN FOR PRESIDENT, DONALD J. TRUMP FOR PRESIDENT, INC.
- 4 stances and 13 issues.



Dataset Details
Weak Label Generation
 Brief Description of the Framework
Analysis of Messaging

#### Weak Label for Stances

- Based on known associations between funding entities and stances, 10% of the funding entities and their ads have stance labels.
  - i.e., Biden Victory Fund, Keep Trump in office.

 Collect 30 news articles covering each issue from left leaning, right leaning, and neutral news media. (<u>https://mediabiasfactcheck.com</u>)

- Collect 30 news articles covering each issue from left leaning, right leaning, and neutral news media. (<u>https://mediabiasfactcheck.com</u>)
- Calculate **Pointwise Mutual Information (PMI)** to identify issue-specific lexicons.

- Collect 30 news articles covering each issue from left leaning, right leaning, and neutral news media. (<u>https://mediabiasfactcheck.com</u>)
- Calculate **Pointwise Mutual Information (PMI)** to identify issue-specific **lexicons**.
- Assign each n-gram to the issue with the highest PMI and build an n-gram lexicon for each issue.

- Collect 30 news articles covering each issue from left leaning, right leaning, and neutral news media. (<u>https://mediabiasfactcheck.com</u>)
- Calculate **Pointwise Mutual Information (PMI)** to identify issue-specific **lexicons**.
- Assign each n-gram to the issue with the highest PMI and build an n-gram lexicon for each issue.
  - Unigrams, resulting in 905 issue-indicating words.



	Dataset Details
	Weak Label Generation
• • •	Brief Description of the Framework
	Analysis of Messaging

## **Graph Embedding Model**

- Embed the following instances in a common embedding space:
  - Ads
  - Funding Entities
  - Issue Lexicon
  - Issue Labels
  - Stance Labels.



## **Graph Embedding Model**

- Embed the following instances in a common embedding space:
  - Ads
  - Funding Entities
  - Issue Lexicon
  - Issue Labels
  - Stance Labels.
- Maximize the similarity between two instances in the embedding space.



## **Graph Embedding Model**

- Embed the following instances in a common embedding space:
  - Ads
  - Funding Entities
  - Issue Lexicon
  - Issue Labels
  - Stance Labels.
- Maximize the similarity between two instances in the embedding space.
- Follow **negative sampling** approach to learn the embeddings.





	Dataset Details
	Weak Label Generation
• · · · ·	Brief Description of the Framework
	Analysis of Messaging

#### Are messages distinctive in ads?

Pro-biden		Anti-trump		Pro-trump		Anti-biden	
Trigrams	P_Cor	Trigrams	P_Cor	Trigrams	P_Cor	Trigrams	P_Cor
vote joe biden	0.763	defeat donald trump	0.594	president trump need	0.667	joe biden democratic	0.713
presidential election held	0.470	request ballot today	0.526	trump need vote	0.729	dont let america	0.357
joe kamala democrat	0.579	time running out	0.238	vote november $3^{rd}$	0.539	taken joe biden	0.600
today vote democrat	0.583	not authorized candidate	0.353	3 <sup>rd</sup> president trump	0.510	radical left taken	0.431
vote democrat joe	0.683	affordable care act	0.382	fake news medium	0.303	democratic party dont	0.524
defeat donald trump	0.535	new trumpcare plan	0.572	president trump spent	0.523	million illegal immigrants	0.295
sure joe biden	0.659	health insurance affordable	0.206	poll sleepy joe	0.396	trillions new taxes&amnesty	0.460
endorse joe biden	0.740	save big health	0.199	live american dream	0.381	biden embraced policy	0.575
joe biden president	0.689	defeat trump gop	0.608	forgotten men woman	0.185	policy far left	0.466
kamala democrat country	0.407	condemn donald trump	0.446	equal opportunity justice	0.206	reduction police funding	0.231

• *pearson\_corr*( $t_i$ ,  $s_j$ ), where  $t_i$  = trigrams generated from  $s_j$  and  $s_j$  = ads related to each stance category

Pro-biden		Anti-trump		Pro-trump		Anti-biden	
Trigrams	P_Cor	Trigrams	P_Cor	Trigrams	P_Cor	Trigrams	P_Cor
vote joe biden	0.763	defeat donald trump	0.594	president trump need	0.667	joe biden democratic	0.713
presidential election held	0.470	request ballot today	0.526	trump need vote	0.729	dont let america	0.357
joe kamala democrat	0.579	time running out	0.238	vote november $3^{rd}$	0.539	taken joe biden	0.600
today vote democrat	0.583	not authorized candidate	0.353	3 <sup>rd</sup> president trump	0.510	radical left taken	0.431
vote democrat joe	0.683	affordable care act	0.382	fake news medium	0.303	democratic party dont	0.524
defeat donald trump	0.535	new trumpcare plan	0.572	president trump spent	0.523	million illegal immigrants	0.295
sure joe biden	0.659	health insurance affordable	0.206	poll sleepy joe	0.396	trillions new taxes&amnesty	0.460
endorse joe biden	0.740	save big health	0.199	live american dream	0.381	biden embraced policy	0.575
joe biden president	0.689	defeat trump gop	0.608	forgotten men woman	0.185	policy far left	0.466
kamala democrat country	0.407	condemn donald trump	0.446	equal opportunity justice	0.206	reduction police funding	0.231

• Pro-biden: vote joe biden, today vote democrat, endorse joe biden.

Pro-biden		Anti-trump		Pro-trump		Anti-biden	
Trigrams	P_Cor	Trigrams	P_Cor	Trigrams	P_Cor	Trigrams	P_Cor
vote joe biden	0.763	defeat donald trump	0.594	president trump need	0.667	joe biden democratic	0.713
presidential election held	0.470	request ballot today	0.526	trump need vote	0.729	dont let america	0.357
joe kamala democrat	0.579	time running out	0.238	vote november 3 <sup>rd</sup>	0.539	taken joe biden	0.600
today vote democrat	0.583	not authorized candidate	0.353	3 <sup>rd</sup> president trump	0.510	radical left taken	0.431
vote democrat joe	0.683	affordable care act	0.382	fake news medium	0.303	democratic party dont	0.524
defeat donald trump	0.535	new trumpcare plan	0.572	president trump spent	0.523	million illegal immigrants	0.295
sure joe biden	0.659	health insurance affordable	0.206	poll sleepy joe	0.396	trillions new taxes&amnesty	0.460
endorse joe biden	0.740	save big health	0.199	live american dream	0.381	biden embraced policy	0.575
joe biden president	0.689	defeat trump gop	0.608	forgotten men woman	0.185	policy far left	0.466
kamala democrat country	0.407	condemn donald trump	0.446	equal opportunity justice	0.206	reduction police funding	0.231

- Pro-biden: vote joe biden, today vote democrat, endorse joe biden.
- Anti-trump: defeat donald trump, defeat trump gop, new trumpcare plan.

Pro-biden		Anti-trump		Pro-trump		Anti-biden	
Trigrams	P_Cor	Trigrams	P_Cor	Trigrams	P_Cor	Trigrams	P_Cor
vote joe biden	0.763	defeat donald trump	0.594	president trump need	0.667	joe biden democratic	0.713
presidential election held	0.470	request ballot today	0.526	trump need vote	0.729	dont let america	0.357
joe kamala democrat	0.579	time running out	0.238	vote november $3^{rd}$	0.539	taken joe biden	0.600
today vote democrat	0.583	not authorized candidate	0.353	3 <sup>rd</sup> president trump	0.510	radical left taken	0.431
vote democrat joe	0.683	affordable care act	0.382	fake news medium	0.303	democratic party dont	0.524
defeat donald trump	0.535	new trumpcare plan	0.572	president trump spent	0.523	million illegal immigrants	0.295
sure joe biden	0.659	health insurance affordable	0.206	poll sleepy joe	0.396	trillions new taxes&amnesty	0.460
endorse joe biden	0.740	save big health	0.199	live american dream	0.381	biden embraced policy	0.575
joe biden president	0.689	defeat trump gop	0.608	forgotten men woman	0.185	policy far left	0.466
kamala democrat country	0.407	condemn donald trump	0.446	equal opportunity justice	0.206	reduction police funding	0.231

- Pro-biden: vote joe biden, today vote democrat, endorse joe biden.
- Anti-trump: defeat donald trump, defeat trump gop, new trumpcare plan.
- Pro-trump: president trump need, vote november 3<sup>rd</sup>, live american dream.

Pro-biden		Anti-trump Pro-trump		Anti-biden			
Trigrams	P_Cor	Trigrams	P_Cor	Trigrams	P_Cor	Trigrams	P_Cor
vote joe biden presidential election held	0.763 0.470	defeat donald trump request ballot today	0.594 0.526	president trump need trump need vote	0.667 0.729	joe biden democratic dont let america	0.713 0.357
joe kamala democrat	0.579	time running out	0.238	vote november 3rd	0.539	taken joe biden	0.600
today vote democrat	0.583	not authorized candidate	0.353	3 <sup>rd</sup> president trump	0.510	radical left taken	0.431
vote democrat joe	0.683	affordable care act	0.382	fake news medium	0.303	democratic party dont	0.524
defeat donald trump	0.535	new trumpcare plan	0.572	president trump spent	0.523	million illegal immigrants	0.295
sure joe biden	0.659	health insurance affordable	0.206	poll sleepy joe	0.396	trillions new taxes&amnesty	0.460
endorse joe biden	0.740	save big health	0.199	live american dream	0.381	biden embraced policy	0.575
joe biden president	0.689	defeat trump gop	0.608	forgotten men woman	0.185	policy far left	0.466
kamala democrat country	0.407	condemn donald trump	0.446	equal opportunity justice	0.206	reduction police funding	0.231

- Pro-biden: vote joe biden, today vote democrat, endorse joe biden.
- Anti-trump: defeat donald trump, defeat trump gop, new trumpcare plan.
- Pro-trump: president trump need, vote november 3<sup>rd</sup>, live american dream.
- Anti-biden: radical left taken, million illegal immigrants, reduction police funding.

• *pearson\_corr*( $t_{i,k}$ , $s_{j,k}$ ), where  $t_{i,k}$  = trigrams generated from  $s_j$  on k issue and  $s_j$  = ads related to each stance category on k issue.





(b) trigrams from anti-trump covid ads



(d) trigrams from anti-biden covid ads

0.48

anti-trump

pro-trump

trump reopen country -

meet american patriot -

-0.2

-0.4

0.35

0.3

anti-biden

- 0.6

- *pearson\_corr*( $t_{i,k}$ , $s_{j,k}$ ), where  $t_{i,k}$  = trigrams generated from  $s_j$  on k issue and  $s_j$  = ads related to each stance category on k issue.
- Pro-biden: Biden's plan for tackling coronavirus by putting effort to develop vaccine, rebuilding country.





- *pearson\_corr*( $t_{i,k}$ , $s_{j,k}$ ), where  $t_{i,k}$  = trigrams generated from  $s_j$  on k issue and  $s_j$  = ads related to each stance category on k issue.
- Pro-biden: Biden's plan for tackling coronavirus by putting effort to develop vaccine, rebuilding country.
- Anti-trump: Rising covid cases and deaths in Florida and Trump's ignorance regarding health experts.





- *pearson\_corr*( $t_{i,k}$ , $s_{j,k}$ ), where  $t_{i,k}$  = trigrams generated from  $s_j$  on k issue and  $s_j$  = ads related to each stance category on k issue.
- Pro-biden: Biden's plan for tackling coronavirus by putting effort to develop vaccine, rebuilding country.
- Anti-trump: Rising covid cases and deaths in Florida and Trump's ignorance regarding health experts.
- Pro-trump: Reopening country.



- *pearson\_corr*( $t_{i,k}$ , $s_{j,k}$ ), where  $t_{i,k}$  = trigrams generated from  $s_j$  on k issue and  $s_j$  = ads related to each stance category on k issue.
- Pro-biden: Biden's plan for tackling coronavirus by putting effort to develop vaccine, rebuilding country.
- Anti-trump: Rising covid cases and deaths in Florida and Trump's ignorance regarding health experts.
- Pro-trump: Reopening country.
- Anti-biden: Unemployment benefit stimulus.





- Targeted demographics by the advertisers:
  - Pro-biden and anti-trump ads mostly target 35 54 age group and the female gender.



- Targeted demographics by the advertisers:
  - Pro-biden and anti-trump ads mostly target 35 54 age group and the female gender.
  - Pro-trump and anti-biden ads target the 55 64 age group and the male population mostly.



- Ad impressions by the demographics:
  - Pro-biden and anti-trump ads are viewed mainly by female audiences and it's highly statistically significant (p-value < 0.01).</p>



- Ad impressions by the demographics:
  - Pro-biden and anti-trump ads are viewed mainly by female audiences and it's highly statistically significant (p-value < 0.01).</li>
  - Pro-trump and anti-biden ads are mostly viewed by males from age range 18 54 (p-value < 0.05)</li>





Significant F-test for the hypothesis of sum of average poll count of each presidential candidate for each day *Granger causes* ad impressions (p-value < 0.05).</li>



- Significant F-test for the hypothesis of sum of average poll count of each presidential candidate for each day *Granger causes* ad impressions (p-value < 0.05).</li>
- No significant Granger causality from ad impressions to polls.



- Significant F-test for the hypothesis candidate for each day Granger cal
- No significant Granger causality from

When polls lean towards one candidate, Facebook ads sponsored by the advertisers supporting that candidate get more attention.

 Formulate a novel problem of exploiting weak supervision to analyze the landscape of political advertising on social media.

- Formulate a novel problem of exploiting **weak supervision** to analyze the **landscape of political advertising** on social media.
- Propose a weakly supervised graph embedding based framework to identify political stance of advertisers as well as the published content and issues of the content.

- Formulate a novel problem of exploiting weak supervision to analyze the landscape of political advertising on social media.
- Propose a weakly supervised graph embedding based framework to identify political stance of advertisers as well as the published content and issues of the content.
- Conduct **quantitative** and **qualitative** analysis on real world dataset to demonstrate the effectiveness of our proposed model.

- Formulate a novel problem of exploiting weak supervision to analyze the landscape of political advertising on social media.
- Propose a weakly supervised graph embedding based framework to identify political stance of advertisers as well as the published content and issues of the content.
- Conduct **quantitative** and **qualitative** analysis on real world dataset to demonstrate the effectiveness of our proposed model.
- Our code and dataset are publicly available at <u>https://github.com/tunazislam/weaklysup-FB-ad-political</u>

#### THANK YOU ③

Slide: <u>https://tunazislam.github.io/files/ICWSM2023\_PoliticalFBAd.pdf</u>

#### Questions?

#### Tunazzina Islam

Department of Computer Science, Purdue University, West Lafayette, IN. Email: islam32@purdue.edu



https://tunazislam.github.io/







#### **Backup Slides**

#### Results

- Stance prediction:
  - Accuracy: 73.71%
  - Macro-average F1 score: 62.81%



#### Results

- Stance prediction:
  - Accuracy: 73.71%
  - Macro-average F1 score: 62.81%
- Issue prediction:
  - Accuracy: 68.39%
  - Macro-average F1 score: 54.76%

