

11-830 Computational Ethics for NLP

Lecture 12: Computational Propaganda



Carnegie Mellon University

Language Technologies Institute

History of Propaganda

- Carthago delenda est!



History of Propaganda

- Carthago delenda est!
- History is written by the winners
 - So its biased, (those losers never deserved to win anyway)
- Propaganda has exists even before writing
- But with mass media its become more refined
 - Newspapers/pamphlets
 - Radio/Movies/TV/News
 - Social Media
 - Interactive Social Media (comments)
 - Personalized Propaganda targeted specially to you sitting quietly in the second row

Propaganda vs Persuasion

- Propaganda is designed to influence people **emotionally**
- Persuasion is designed to influence people with **rational arguments** (ish)

- But its not that easy to draw the line **objectively**
 - They use propaganda to influence
 - We use rational arguments to inform

We vs Them

We have ... Army, navy and air force Reporting guidelines Press briefings	They have ... A war machine Censorship Propaganda
We ... Take out Suppress Dig in	They ... Destroy Kill Cower in their fox holes
Our men are ... Boys Lads	Their men are ... Troops Hordes

The Guardian 1990

Propaganda

- Demonize the enemy
 - “The only good bug is a dead bug”
- Personalize your side
 - “Our good boys ...”
- Be inclusive
 - “Good people like yourself ...”
- Be exclusive
 - “Never met a good one ...”



Propaganda

- Obfuscate the source
- Nazi Germany makes a BBC-like show
 - Lord Haw Haw (William Joyce) “Germany Calling”
 - Sounded like a BBC broadcast (at first)
 - Talked about failing Allied Forces
 - Personalized to local places
- Flood with misinformation
 - To hide main message
 - Discredit a legitimate source
 - Add a sex story to deflect attention

Propaganda

- Doesn't need to True (or False)
 - Make up stories that distract
- But you can still just be selective with the truth
 - Marketing does this all the time
 - The most popular smart phone in the world
 - The most popular smart phone platform in the world
- Maybe truth plus distraction
 - Add a hint of a financial scandal

Public Relations Office

- Most countries, organizations, companies have official press releases
 - Mostly legitimate news stories
 - But may sometimes just propaganda
 - The mixture with legitimate news strengthens the illegitimate
- Major News Outlets have explicit bias
 - VOA, RT, Al Jazeera, BBC World Service, DW
- Private News Organizations have explicit bias
 - Washington Post (owned by Jeff Bezos)
 - Blog sites (owned by unexpected rival)
 - Often explicit bias statement

Computational Propaganda

- People still generate base stories
- But automated bots can magnify attention
 - Bots can retweet
 - Add likes
 - Give a quote and a link
- Build an army of bot personas
 - Be applied to many aspects of on-line influence

Computational Propaganda Project

University of Oxford

- Philip N Howard and Sam Woolley
- Since 2012
- Originally at University Washington (started with an NSF grant)
- Grants on
 - Computational Propaganda
 - Misinformation, Media and Science
 - The Production and Detection of Bots
 - Restoring Trust in Social Media Civic Engagement
- They produce (detailed) reports on aspects of
 - Fake News, Election Rigging
 - Regulation of Social Media

Political Bots

- @Girl4TrumpUSA created on Twitter
- Generated 1,000 tweets a day
- Mostly posting comments and links to Russian news site
- Deleted by Twitter after 38,000 tweets
- Many other similar bots
 - They amplify a candidate's support
 - Forward other messages (so you see things multiple times)
 - Ask: "what do you think about 'x'" (to get responses)
 - Like and retweet articles
 - Create fake trends on hastags
 - Astroturfing vs grass roots
 - Manufacture consent

How Many Bots

- Use crowd sourcing services to do tasks
- Can buy armies of bots with existing personas
- Start a twitter account
 - Buy a following of bots
 - High number followers attracts real followers
 - Bots will get deleted
 - Keep all the real followers
- There are offers of 30,000 personas for sale



Bot Detection

- Not very hard (at present)
 - Bot activity over time is quite different from humans
 - Bot post contents is often formulaic (its all rule driven)
- Oxford Computational Propaganda Project
 - Published papers on bot types and detection techniques
 - They interviewed a bot maker
 - “How do you avoid your bots from being detected”
 - “We read papers by you on what you do to detect us”
- Oxford Computational Propaganda Project
 - Looking for post doc to work on bot detection



Bot Development

- Bot content formulaic
 - Generated from basic templates
 - Hand written
- Bot actions vs machine learning
 - Reinforcement learning
 - Send message1 to 50 people
 - Send message2 to different 50 people
 - Count number of clicks
 - Send most clicked message to 500 people
- Do this on more targeted messages to personalized interests
 - Send education message to person who mentioned education
 - Send healthcare message to person who mentioned healthcare

Automated Bot plus Humans

- But Crowdworkers wont post propaganda for you
 - So ..
- Please help with this propaganda detection problem
 - Here are 4 messages
 - Which ones are real, and which ones are bot generated:
 - “We’re the greatest”
 - “They’re the worst”
 - “Where is his birth certificate?”
 - “My granddaughter sent this link ...”
- Thank you for help with the propaganda generation problem

Investigative Journalism on Bots

- FCC Net Neutrality Public Comments
 - Overwhelmingly anti-neutrality



Investigative Journalism on Bots

- FCC Net Neutrality Public Comments
 - Overwhelmingly anti-neutrality
- Dell Cameron and Jason Prechtel, Gizmodo
 - Traced each comment (uploaded through API)
 - Traced timing with downstream registrations
 - Highly correlated with PR firms CQ Roll Call and Center for Individual Freedom (CFIF).
- “Of course there are fake posts to the comment section, see look we added a few on the other side too so we don’t look so shady.” (not a quote)

Is it all bad Propaganda

- Probably we can't draw the line between propaganda and persuasion
-
- Social media use for protests can be effective
 - 4Chan/Anonymous and the Arab Spring 2010/11
 - Soc.culture.china (usenet) and Tiananmen Square Protests 1989
- Much of early Internet Interest was in the voice of the people
 - Cyberactivists (John Perry Barlow, John Gilmore) saw social media as a plus
 - “A Declaration of Independence of Cyberspace”
 - Electronic Frontier Foundation

Comparison to Spam

- Spam: the mass distribution of ads (real or otherwise)
- It was successful at first (a few people clicked)
- People developed automatic spam detection algorithms
 - Mostly on usenet as that was the largest forums at the time
 - Then in email
 - Detection improved, but its still there
- We still receive spam, though mostly we ignore it
- Other much more sophisticated marketing is now common
 - And more acceptable
 - Google links to purchasing options
 - Amazon recommendations
- So spam is contained and mostly ignored

Can Propaganda become like Spam

- People send spam **if** it works
 - Spam working means people “buying”
- People send propaganda **if** it works
 - Propaganda working means people ... voting (?)
 - Which isn't as important as buying the best smart phone :-)
- People may become more sophisticated with propaganda
 - Learn to ignore it, (but what of those who don't)
 - But it will become more targeted to the unsophisticated
- Propaganda messages may become more sophisticated
 - Control your news bubble/echo chamber
- Propaganda messages may drift to informative messages
 - People will learn to evaluate both sides of the issue and make informed decisions