Research Design Outline

Detecting Bots on Reddit
Computational Social Science

Matteo Mazzarelli

January 27, 2025

1. Research Question, Societal & Scientific Relevance

Research Question: How do bots influence discussions and information sharing within Reddit communities (subreddits)?

Hypotheses:

- Bots amplify specific narratives.
- Bots contribute to polarization/echo chambers.

Relevance:

- **Societal Relevance:** This research is relevant for combating misinformation, maintaining healthy online communities, and understanding public discourse manipulation. As social media platforms become increasingly influential in shaping public opinion, it is crucial to understand the role of automated accounts in these spaces. Bots can be used to spread propaganda, manipulate trends, and create an illusion of consensus, potentially leading to harmful real-world consequences.
- **Scientific Relevance:** This study contributes to the understanding of online social influence, bot detection, and information diffusion on Reddit. It can help refine existing bot detection methods and provide insights into how information, particularly misinformation, spreads within online communities. The findings can inform the development of strategies to mitigate the negative impacts of bots on online discourse.

2. Data & Measurement & Planned Analyses

Data

- **Source:** Reddit API (using the Python Reddit API Wrapper, PRAW)

- **Subreddits:** Selection based on a scoring system that considers both the number of subscribers and the presence of keywords associated with topics likely to be targeted by bots (e.g., politics, elections, news, war, vaccines, conspiracy, finance, etc.). The list of keywords can be generated with the help of a large language model (e.g. Gemini) tuned for reproducibility (stochastic parameters such as Temperature turned to 0) and includes:

```
politics elections government news worldnews
conspiracy war military vaccines health
covid finance crypto stocks technology
ai socialmedia propaganda disinformation activism
```

The top 10 subreddits based on this scoring system are:

	Name	Subscribers	Submission Type	Bot Score
3	worldnews	43565302	link	5
13	news	29223434	link	4
15	videos	26808936	link	4
22	AmItheAsshole	23083122	self	4
23	explainlikeimfive	22977358	self	4
26	Art	22367695	link	4
27	sports	21708101	link	4
28	dataisbeautiful	21402346	link	4
30	personalfinance	20265058	self	4
31	UpliftingNews	20189860	link	4

- **Collection:** Data collected includes posts, comments, user data (if available), and timestamps.

Measurement

- **Bot Identification:** A heuristic approach will be used to identify potential bots. This approach will consider factors such as account age, karma, post frequency, and repetitive content. For instance, accounts with very low karma, extremely high posting frequency, or those exhibiting highly repetitive posting patterns may be flagged as potential bots.
- **Analysis:**
  - **Descriptive:** The prevalence of bots within the selected subreddits will be estimated. Content analysis will be performed to compare the content generated by suspected bots versus human users. Engagement patterns (e.g., upvotes, downvotes, comment frequency) of suspected bots will be analyzed.

– **Inferential:** Correlation analysis will be conducted to examine the relationship between bot activity and various factors, such as the prevalence of specific narratives or the level of polarization within discussions.

3. Challenges & Open Questions

Challenges

- **Difficult Bot Detection on Reddit:** Reddit's relative anonymity and the limited availability of tools compared to other platforms make bot detection challenging.
- **Creating a Reliable "Ground Truth" Dataset:** There is no readily available "ground truth" dataset for bot detection on Reddit. This makes it difficult to train and evaluate bot detection models.
- **Data Access Uncertainties:** The Reddit API has rate limits, which may affect the amount of data that can be collected within a given timeframe.

Discussion

The chosen methods and data have limitations. The heuristic approach to bot detection is not perfect and may result in misclassification. The reliance on the Reddit API introduces potential biases, as the API may not provide a completely representative sample of Reddit activity.

The findings of this study could have implications for Reddit and platform policies. If a significant presence of bots is found to be influencing discussions, Reddit may need to implement stricter measures to detect and remove bots. The findings could also inform broader discussions about platform governance and the responsibility of social media companies in mitigating the spread of misinformation.

Open Questions

- How can bot detection on Reddit be improved, given the platform's unique characteristics?
- What are the most effective strategies to counter the influence of bots on Reddit and other online platforms?

4. Population (Target)

The target population is users of the social media platform Reddit, specifically those participating in subreddits related to topics often targeted by bots, such as news, politics, and finance. This includes both human users and automated accounts (bots).

5. Sample

The sample consists of data collected from the top 10 subreddits identified as most vulnerable to bot influence, based on the scoring system described in the "Data" section. These subreddits are selected based on their large subscriber base and the presence of keywords related to bot-targeted topics in their descriptions.

6. Unit

The fundamental units of analysis are individual posts and comments within the selected subreddits. The analysis will consider both individual users (or suspected bots) as well as aggregate patterns within each subreddit.

7. Data/Observations

The data collected includes posts and comments from the selected subreddits, along with associated metadata such as timestamps, user information (where available), and engagement metrics (upvotes, downvotes). The data is structured as a combination of cross-sectional (snapshot of activity at a given time) and panel data (tracking posts and comments over time). For example, a post's initial reception can be captured (cross-sectional), while changes in engagement over time can also be tracked (panel).

8. Measurement

The different variables are measured as follows:

- **Bot Activity:** Measured through a heuristic approach based on user account data and posting behavior.
- **Narrative Amplification:** Measured by analyzing the frequency and prominence of specific keywords and phrases in posts and comments.
- **Polarization/Echo Chambers:** Measured by analyzing the sentiment and language used in discussions, as well as the network structure of interactions between users (e.g., identifying clusters of users who primarily interact with each other).

9. Machine Learning Method

While the primary focus is not on building a machine learning model, descriptive and inferential analysis will be conducted. These could include:

- **Descriptive Statistics:** Calculating the prevalence of suspected bots, analyzing the distribution of post and comment lengths, and examining engagement patterns.
- **Correlation Analysis:** Investigating the relationship between bot activity and other variables, such as the frequency of specific narratives or the level of polarization.

10. ML Model & Unit of Analysis & Estimation

As the main analysis is not focused on machine learning, this section is not applicable.

11. Theory/Mechanisms

The underlying theory is that bots can be used to manipulate online discussions by amplifying specific narratives, creating echo chambers, and influencing public opinion. This is achieved through various mechanisms, such as:

- **Agenda Setting:** Bots can artificially inflate the visibility of certain topics by posting frequently and generating engagement (e.g., upvotes, comments).
- **Framing:** Bots can promote specific interpretations of events or issues by using particular language and framing techniques.
- **Social Proof:** Bots can create the illusion of consensus or popularity by mimicking human behavior and interacting with other users.

12. Training/Testing

Not applicable as the main focus is not on building a machine learning model.

13. Assessment of Accuracy

Not applicable.

14. Previous Studies

Previous research on bot detection and influence on social media will be reviewed. This includes studies on platforms like Twitter and Facebook, as well as research on Reddit. The literature review will focus on identifying different bot detection methods, analyzing the impact of bots on online discussions, and understanding the mechanisms through which bots exert influence. This will help inform the current study's methodology and provide a context for interpreting the findings.