June 4, 2021

Group I

Vincent Wahid Ajoubi	viaj@itu.dk	BSc Data Science
Edi Begovic	edbe@itu.dk	BSc Data Science
Høgni Jacobsen	hoja@itu.dk	BSc Data Science
Mourad Baptiste Karib	mouk@itu.dk	BSc Data Science

Reflections on Data Science 2021

Social influence: Reddit upvotes

Introduction

Reddit is the place where thousands of ideas are shared daily. Whenever a post is created, by default the platform gives it an upvote, where the number of upvotes determines the rank of the post. The higher the rank, the more popular the idea or comment of the post is. Users themselves decide the popularity of the idea afterwards by up/downvoting. In this experiment we will be conducting a test for the presence of the *Matthew effect* [1](“rich get richer”) through a randomized control trial (RCT).
We want to evaluate the impact of an initial upvote for random posts shared on the platform regardless of content. We split the posts into two groups, one that gets the treatment and a control, with the treatment group receiving an additional initial upvote from us and the control group left untouched by us.

Methodology

200 new posts (from the r/all [2] subreddit) were scraped each day for a week, all with an initial upvote count of 1 and with 0 comments - the defaults for new posts on Reddit. Half of the posts (100) were assigned to the treatment group and their posts were given an upvote, while the other half were left untouched as a control. For each day subsequent to the post being scraped, and for a total of 6 days, we noted the latest upvote count and the number of

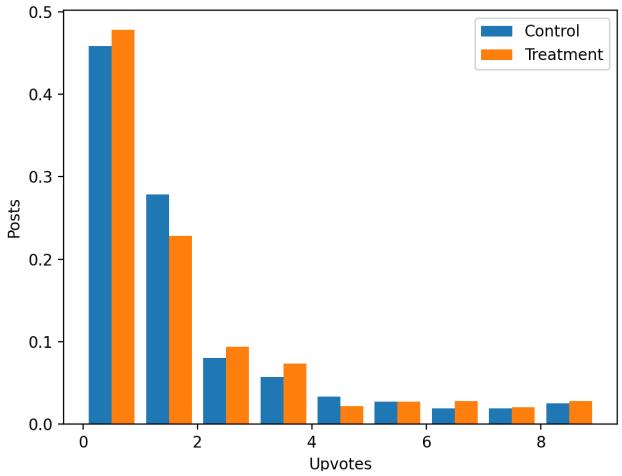
comments on each post. That is, posts pulled on day 1 would be completed by day 7, while posts pulled on day 7 were completed by day 14.
This was implemented using Praw [3] for Python (API Wrapper) with access to the Reddit API via a developer account.

To assess the influence of the treatment on the two groups, we define two outcomes: either *successful* or *unsuccessful*. Posts with a score (post upvotes) higher than 10 points are deemed *successful*. With this outcome, we simply perform a chi-square test on the proportions of successful/unsuccessful posts for each group and check if they differ significantly.

While we could have performed a one-sample T-test on the mean scores of the two groups, it would have been very sensitive to outliers as post scores follow a power law. This could partially have been mitigated by instead using log scores, thus reducing the impact of viral posts.

Results

We found that the score stabilized quite rapidly. Comparing the scores of posts (posts with scores between 0-100) on day 3 and 7, we found that 92% of post scores remained unchanged on day 7, with 99% being within 2 points.



From the above histogram of post scores, it's clear the distribution is logarithmic, affirming the pareto principle (*like distribution of wealth in a society*). Close to ⅔ of the collected posts had less than 3 points, while we have less posts with more than 6 up-votes. Given these properties, we decided to deem a post a "success" if it had a score or 10 or more.

We found that both the mean scores of the treatment group as well as the number of ‘successful posts’ (posts with a score greater than 10) were smaller than the control group. Thus even if the relationship is significant, we can already conclude that there is no *positive* impact of an initial upvote to a post's chances of spreading. After performing a chi-square test on the proportions of successful posts, we found no significant relationship between the treatment and the outcome with a p-value of 0.053.

Discussion

The number of users who get to see a post as part of their home feed is determined by Reddit's (undisclosed) algorithm. We therefore cannot know what impact other factors such as account age, post history, and general activity may have on the weight of an individual user’s vote. For instance, the weight of a user's vote may be discounted relative to how many posts they’ve previously voted on and in what timespan they were voted.

It is known that Reddit has implemented mechanisms to reduce voting manipulation through *shadow banning* [4]. Shadow banning is a way of deterring bots or malicious people from misusing a service by hiding their posts and other activity from other users *apart* from themselves. That way, it’s harder for an adversary to assess the effectiveness of their work, as things appear normal from their perspective.

Another related factor is the impact of a single upvote. There may be a cumulative advantage for posts to get an initial boost to their score, but is only significant when the treatment is above a certain threshold.

Conclusion · Future work

The treatment we applied wasn’t sufficient to influentiate up-votes on Reddit. Either we should use various accounts to increase the number of up-votes or consider generating comments. It is also possible that Reddit’s algorithm has a more complex way to influence a post's exhibition. It would require more research and tests to understand how to hack it, in order to boost a post. It is probably dependent on the profile who up-votes/comments (their information or reputation on the platform). We also know that Reddit is aware of this problem and has algorithms to tackle these problems, implementing shadow-banning and other processes to regulate their page.

Sources

[1] Mathew effect
[2] Reddit r/all
[3] Praw for Python
[4] *An unofficial guide on how to avoid being shadowbanned*

All source code for this project is contained the the following git repository:
https://github.com/edibegovic/ReflectionsDS