Precog Research Paper Report

Paper name : How Community Feedback Shapes User Behavior

Authors : Justin Cheng, Cristian Danescu-Niculescu-Mizil, Jure Leskovec

This research paper delves into the cognitive effect that feedback and rating mechanisms on prevalent social media platforms have on authors of social media posts. Behavioral psychology has introduced the concept of an operant conditioning framework, wherein positive feedback results in a 'reward' stimulus and negative feedback results in a 'punishment' stimulus. Ideally this would imply that authors of badly written content would receive negative feedback and would improve further content or reduce the amount of content they write, and authors of good content would increase their frequency of posting and continue to write good content.

To verify this, the authors of the paper have come up with a methodology to study the behaviour of community feedback by observing posts in 4 communities based on upvotes and downvotes.

One issue faced was the metric to decide how the upvotes were to be taken into account. Each method appeared to have some issues. To decide which of these methods was best, the authors conducted an experiment where they took a sample of people, gave them a set of upvote and downvote pairs, and asked them how they would feel upon receiving that particular pair of upvotes and downvotes. This method showed that the ratio of positive votes to total votes was the best metric to deal with. Personally I would have identified existing users on these social media platforms that have already authored posts of such a situation. These people would be more reliable to get information from as they have physically experienced these stimuli instead of people that only imagined what it would be like.

One more thing the authors needed was a way to measure the quality of a certain post. To that end, they have trained a binomial regression model(output has a binomial distribution) to predict the quality of a post. For training, only the text was taken since they wanted to predict only the quality of the writing. To validate this model, the authors used humans to manually verify the quality of the text. Due to the limits of manual work, the sample size of posts taken was quite small (n=171). Personally I think that a better method would be to take posts that have a high upvote ratio and *which is the first post of the author.* This would remove the community bias and we could safely say that the post was well received for the post quality alone.

To observe the effects of community feedback, the authors took two posts which were of similar quality but got different evaluations. The posts of the same authors were then compared

after these evaluations to determine the effect that they had on the content of these users. Upon observation, contrary to our expectations, we find that positive feedback has little to no effect on the future posts of the user, whereas a negative feedback causes not only worse quality posts from the user, but more frequent posts from the user.

The authors also looked into how community networks were organised in social media. Using structural balance theory and examining the fraction of balanced triads, the authors were able to surmise that the network is most polarised when the number of upvotes and downvotes are split evenly and that the users mostly voted within their own groups.

In terms of further work, one thing that could be improved upon is the accuracy of the model used to predict quality. Improving the model may lead to even better and more detailed relations being discovered within communities.

Other future work may include research on community feedback not only based on rating and upvotes, but on actual textual feedback provided in the comments of posts.

One factor that was not taken into consideration in this paper was the type of person giving feedback. This is a major factor since a more important person would have a higher impact upon a person.

All of these advancements would likely require in depth analysis of the linguistics and the context provided in the post or thread and would probably be extremely interesting and insightful to look into, despite the possible challenges that they pose.