TEAM: DEEP LEARNERS

NIKHIL VASUDEVA RAVI TEJA SUNKARA

I. PAPER

Community Interaction and Conflict on the Web by Srijan Kumar, William L. Hamilton, Jure Leskovec, Dan Jurafsky

Community interaction plays a major role in offline communities and online communities. We interact with people with different opinions and ethnicities and, consequently, misunderstandings are bound to happen. But some people try to use this misunderstanding to further divide people by using strong language and behavior like trolling, sockpuppetry and public shaming. This paper tries to understand what groups of people instigate such conflicts in the online communities of Reddit. What type of defense is good to overcome this conflict? What type of statements creates neutral and negative mobilizations of people? Finding and understanding answers to these questions helps us prevent such conflicts in the future and to understand this is the main motivation to implement this paper.

The paper is complex with various findings using statistical measures and a deep learning model which further motivated us to choose this.

II. SUMMARY

Communities are a major part of online platforms. The authors try to understand the interaction between communities and how they impact users especially conflicts (similar topics but different views). Measures were suggested to mitigate such conflict and a deep learning model was built to predict if a post will create conflict.

A. Social Science Perspective

1. Hypotheses

- Assuming that only posts with hyperlinks to other groups lead to conflicts.
- Considering only pairs of communities
- The opinions and beliefs shared on an anonymous website like Reddit are true and applicable across other web platforms

2. Theoretical Approach

Analyzing users across 36,000 communities by using “cross-links” posted by users to understand the effect it has on the source community (the community in which the hyperlink is

posted) and the target community (the community to which the hyperlink refers to). Carefully control for baseline rates of user activity and trace out various trends.

3. Results

- A small number of communities initiate most conflicts
- Topically similar, but opposing ideology, groups fight
- Conflicts are initiated by active users but are carried out by less active users
- Formation of echo-chambers
- Long-term adverse effects of mobilizations in terms of colonization and immigration
- Ways to defend against conflicts

B. Computational Perspective

Socially-primed LSTM model that combines graph embeddings, user, community, and text Features to predict if a post will create conflict was novel and gave good results.

III. DETAILS OF OUR IMPLEMENTATION

A. Data

40 months of Reddit comments and posts from which embeddings were created by authors. We will use the preprocessed data and embedding provided by the authors. We would also like to explore the data and try to see if the findings are accurate. This would allow us to see some details that the authors might have swept under the rug. We would like to call out that this is a laborious task and we do not know the scale of the problem. We will try our best to get our hands dirty.

B. Tools/Methods

We will be using Pytorch to implement the LSTM model. Python will be our go-to tool to perform statistical tests and trace out the inferences as obtained by authors. Random forest model will be implemented using scikit-learn package in python. We will be replicating

C. Metrics

AUC will be used to test the predictive performances of our models. We will try to match the p-values, standard deviations, and probabilities obtained in the paper.

D. Risks and Limitations

The usage of statistical tests to emulate the results obtained by authors may prove difficult as we are not well-versed with statistics and finding the relations by carefully

controlling for other factors. The page ranking described in the paper to quantify the impact even though we understand we may not be able to implement them.

In the worst-case scenario, we will implement the deep learning model for prediction tasks and improve it by adding attention layers, skip connections and transfer learning. Successfully obtain results related to ‘social hypotheses’ which are not related to page ranking - Fig 3, Fig 5, Fig 7. Improving the ability to predict is by and itself useful as the deep learning models are evolving at a rapid rate. Gaining the understanding of data using statistics and showing the results will validate the authors’ claims and proves that the model can be used for other web interfaces.

We are looking forward to asking and learning wherever necessary as the paper covers a broad range of topics that are useful to learn and understand.

IV. TASK BREAKDOWN

We chose the project as it is not straight-forward and difficult to replicate in most parts. Even the LSTM model uses graphical embeddings and other features that we are not completely aware of. The goal of choosing this project is to challenge ourselves and learn. So there won’t be any individual task breakdown as we will be contributing equally for good or bad.

V. OUR CONTRIBUTION

We are thinking of improving the LSTM model by using the attention layer, skip connections and transfer-learning methods used such as chain-thaw we saw in Felbo et al. paper. To see the effect of these new additions we can compare accuracy on the same data and also if possible use new Redditt data to check how well our model can be transferable. Validating the transfer learning approach is going to be difficult and is most likely not going to be done.