Quantifying Textual and Community-based Factors for Internet Altruism

Bilguunzaya Battogtokh, Clare Zhu

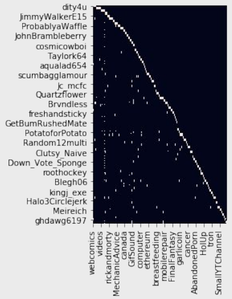
Department of Computer Science, Stanford University

Introduction

- Altruism is a social behavior in living beings that compels individuals to give to others, with no expected reward in return
- Although normally limited in impact to one's immediate friends and family, with the advent of crowdfunding and mass donation campaigns, we have seen **an unprecedented rise in altruism to online strangers**
- Jurafsky et al. used data from Reddit's RandomActsOfPizza to explain the importance of narrative, in addition to classic hypotheses in psychology, to individual altruism
- We improve upon their conclusions using a more nuanced **model of request language** and a greater understanding of the **network ecosystem of givers and receivers**

Dataset

- We use the Reddit Pizza Request dataset provided by Jurafsky et al., which we extend by collecting data using Pushshift.io
- Original dataset contains 6 identified givers' subreddit participation; using the extended dataset, we find **203** additional examples
- In the original dataset (collected in 2014), **25%** of requesters received pizza
- In the extended (more recent) dataset, we find that **58%** of users receive a pizza on average
- This indicates increased quality moderation and spam detection in /r/RAOP



Adjacency matrix between users and subreddits

Approach

```
Title: [Request] My wife has decided that we're going vegan
Body:  I haven't eaten meat in days and I no longer feel manly.
number_of_downvotes_of_request_at_retrieval: 13
number_of_upvotes_of_request_at_retrieval: 20
request_number_of_comments_at_retrieval: 13
requester_received_pizza: False

Title: (REQUEST) Unemployed and hungry in Tucson, AZ
Body:  Unemployed and broke future soldier and his wife. Out of food and could
       use some help, will gladly pay it forward come January when I ship. Thank you.
...
requester_received_pizza: True
```

The task: From the request text and additional features, predict whether the requester receives a pizza.

Left: Examples of requester text and additional features

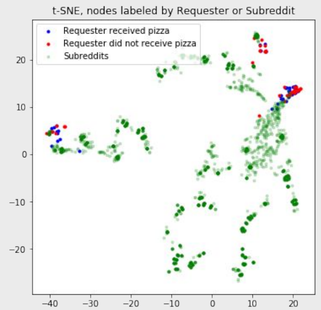
(Pizza from http://www.pngmart.com/image/5622)

Language Modeling

- **Word2Vec** is a 2-layer neural network that is trained to reconstruct the linguistic contexts of words
  - CBOW: predicts target words from context words
    - Use GloVe embeddings: 400k vocab, 100d vectors
    - Text representation: average of all word vectors
  - Logistic regression classifier on text representation
- **BERT (**Bidirectional Encoder Representations from Transformers) is a state-of-the art model for pretrained language representations
  - Goal is to represent variable length sentence into a fixed length vector
  - BERT-Base model (12 layers, 12 attention heads, ~110M parameters) outputs
    - Pre-trained on large amount of plain text data in unsupervised manner
  - Deep neural net classifier with 3 hidden layers [64, 32, 16] on BERT encoding

Network Analysis

- Using only data that represents the connections between users based on subreddits that they participate in, can we improve upon the language modeling result?
- **Baseline:** Use subreddit adjacency as a feature class
- **Improved Model:** Apply network analysis techniques to predict whether requesters receive pizza based on the subreddits they participate in and subreddits of givers
  - node2vec: Node Embeddings using Biased Random Walks
  - Link Prediction from known givers to aspiring receivers
- *Next step* - **Combined Language Model + Network Model:** Concatenate the BERT- and network-generated features for a final classification layer



A t-SNE representation of the graph, created using spectral embeddings of the graph adjacency matrix

Results

- Original dataset (2014)

Test Set Results	AUC-ROC	Accuracy
Baseline: n-grams	0.53	0.53
Jurafsky et al. baseline	0.621	--
Human oracle	0.56	0.71
BERT Language Model	0.58	0.75
Word2Vec	0.50	0.75

- Our extended dataset (2019)

Test Set Results	AUC-ROC	Accuracy
Baseline: n-grams	0.56	0.53
Baseline: Naive Adjacency Features	0.484	0.522
Human oracle	TBD	TBD
BERT Language Model	TBD	TBD
Node Embedding Model	TBD	TBD
Word2Vec	0.50	0.58

Conclusion

- Advanced language modeling outperforms the authors' n-grams approach, as well as the human oracle
- Subreddit participation network data is noisier than the language data, but **has potential to augment language features**
- Extending the dataset to present-day gives us insight into **effects of increased moderation**
- **Application:** Moderators can use language modeling to boost high-quality requests

References

[1] E. Fehr and U. Fischbacher, "The nature of human altruism," *Nature*, vol. 425, no. 6960, p. 785, 2003.

[2] T. Althoff, C. Danescu-Niculescu-Mizil, and D. Jurafsky, "How to Ask for a Favor: A Case Study on theSuccess of Altruistic Requests," in *Proceedings of the International Conference on Weblogs and Social Media (ICWSM)*, 2014.