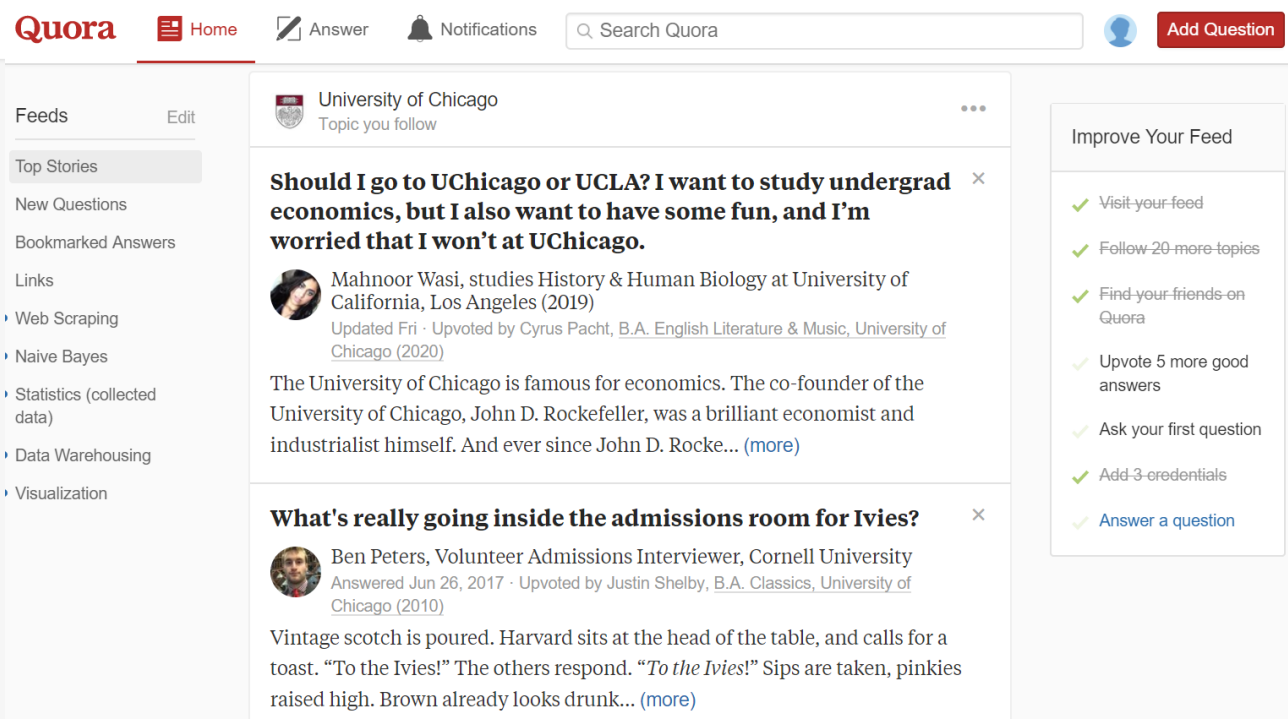
User Behavior in Q&A community: an analysis of Zhihu

Andi Liao

2018/04/04

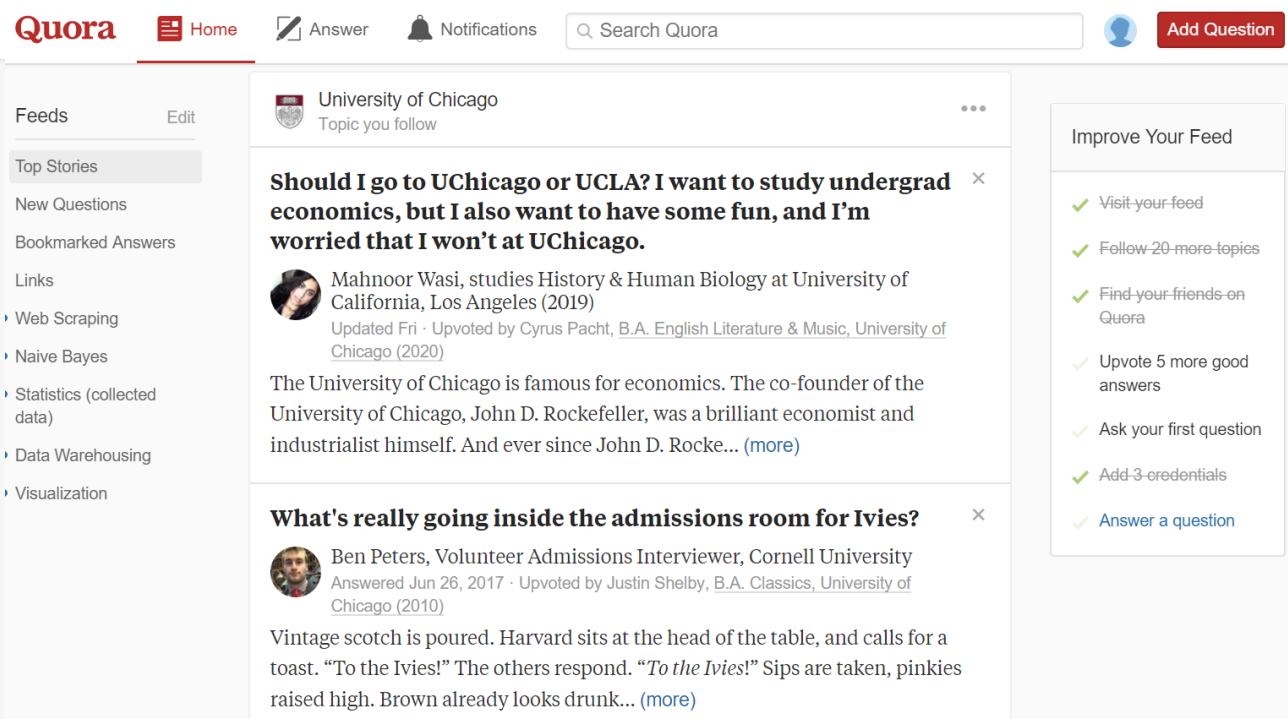
Brief Intro to Zhihu

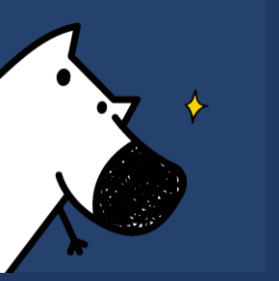
- Similar with Quora
  - Q&A Community - Follower & Following, Upvotes & Downvotes
  - User Interface - Navigate Bar, Topics
  - Recommend System - Feed, Invitation



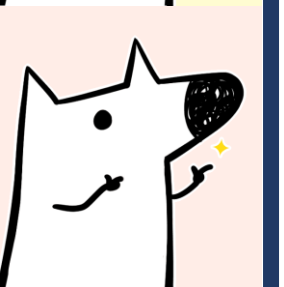
Brief Intro to Zhihu

- Difference with Quora:
  - User Identity - Any Name
  - Upvote - Agree
  - Top Writer - Centered





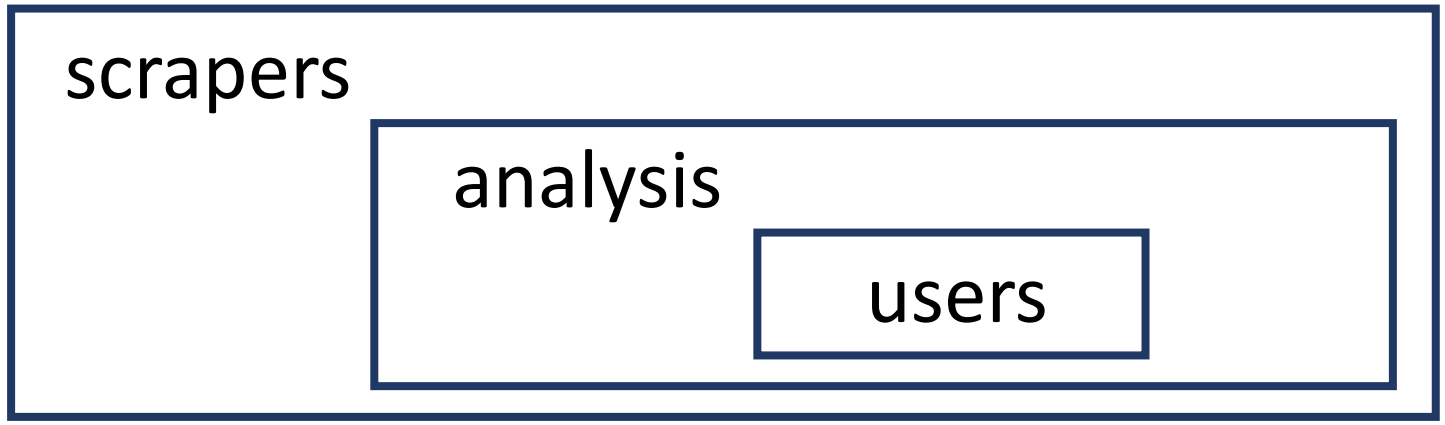
Research Question



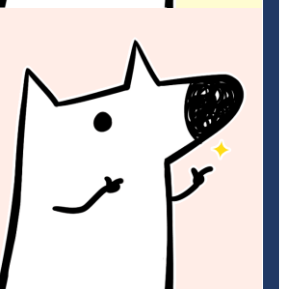
- User Behavior & Interaction in Zhihu Community
  - Clustering:
    - Is it possible to cluster users based on data of their activities?
  - Prediction:
    - Is it possible to predict a new user to be a top writer based on data of his activities?

Why It Matters?

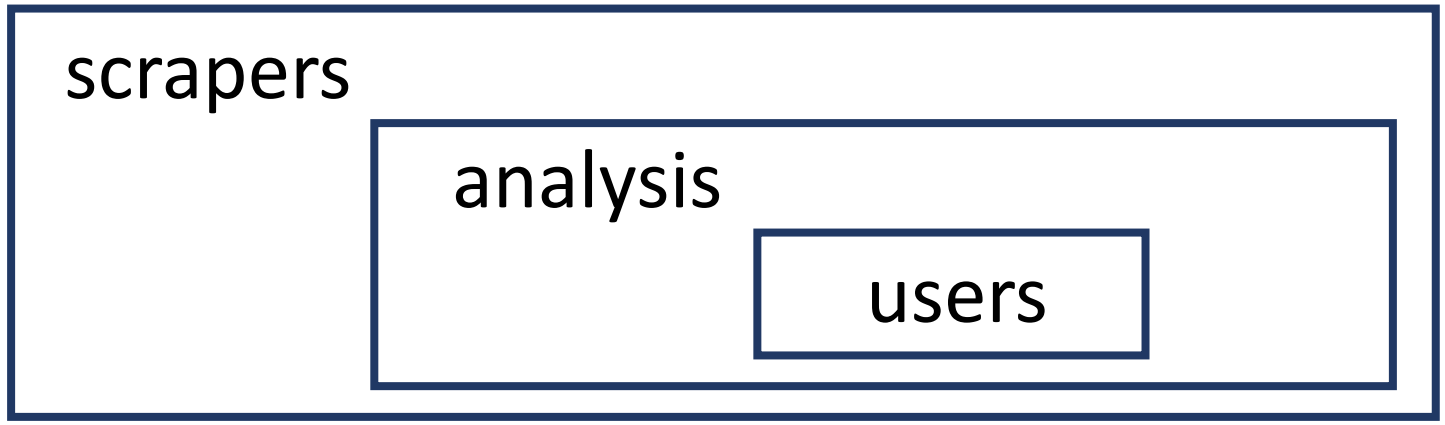
- Mixture: knowledge sharing + social
  - Venmo: Finance + Social
  - Weibo: Nickname + Comment
- Many scrapers, but few analysis focused on users in depth



Why It Matters?



- Mixture: knowledge sharing + social
  - Venmo: Finance + Social
  - Weibo: Nickname + Comment
- Many scrapers, but few analysis focused on users in depth

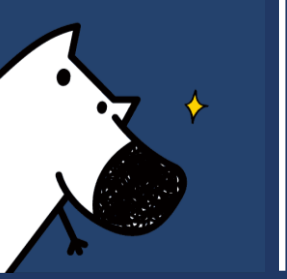


Data Source

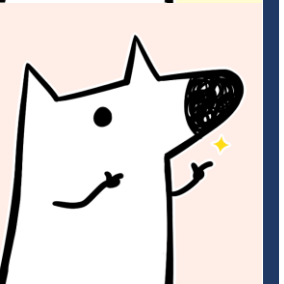
- Scraper: https://github.com/MatrixSeven/ZhihuSpider - Java
  - Backup: https://github.com/7sDream/zhihu-oauth - Python unofficial API

- Data form: Mysql database

Follower	User	User Information
user_name	user_id	company, job
follower_name	index_url	education
update_time	token	answer, question
…	…	…



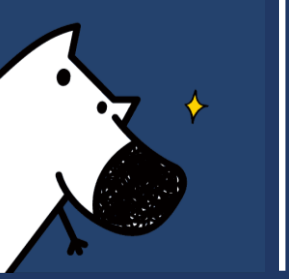
Possible Theory



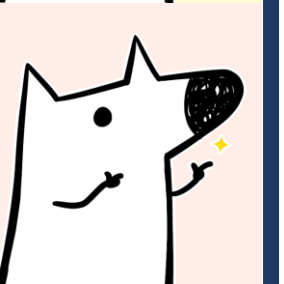
- Social and Interaction Graph
  - Degree-distribution: following, follower
  - Clustering coefficient: similarity
  - Reciprocity and balance: symmetric relationship
  - Assortativity: tend to connect similar nodes in the network
  - Tie Strength: interaction frequency

- Reference:

  - Zhang, X., Tang, S., Zhao, Y., Wang, G., Zheng, H., & Zhao, B. Y. (2017). Cold Hard E-Cash: Friends and Vendors in the Venmo Digital Payments System. In *ICWSM* (pp. 387-396).
  - Wang, T., Chen, Y., Wang, Y., Wang, B., Wang, G., Li, X., ... & Zhao, B. Y. (2016). The power of comments: fostering social interactions in microblog networks. *Frontiers of Computer Science*, *10*(5), 889-907.
  - Wang, G., Gill, K., Mohanlal, M., Zheng, H., & Zhao, B. Y. (2013, May). Wisdom in the social crowd: an analysis of quora. In Proceedings of the 22nd international conference on World Wide Web (pp. 1341-1352). ACM.

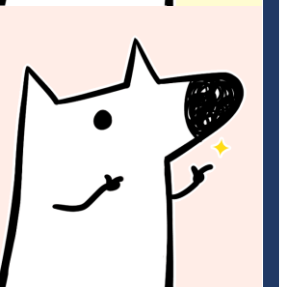


Method & Tool



- Clustering & Prediction
  - K-mean++ clustering
  - Random forest classifiers
- Challenge
  - Supervised learning methods without labels
  - Categorical variables exist
- Reference
  - Patil, S., & Lee, K. (2016). Detecting experts on Quora: by their activity, quality of answers, linguistic characteristics and temporal behaviors. Social network analysis and mining, 6(1), 5.

Expectation



Wikipedia
The Free Encyclopedia

Quora

twitter

stackoverflow

知乎
www.zhihu.com

facebook

Non-Social

Social