Cryptocurrency Recommendations Via Tweets

Perry Chu

Goal

- Relate Tweets to Cryptocurrencies
  - NLP on Tweet text?
  - Recommender tweets or tweeters?

- Recommend new cryptocurrencies based on user interests; infer interest from hashtags
  - Blog: increase views -> $
  - Exchange: increase volume -> $

Workflow

- Data Collection
- Data Processing
- Recommender Model
  - Popularity (Baseline)
  - Collaborative (SVD - similar user)
- Test / Validation

Data

- Crypto-currency tweets
    - API stream tweets with relevant hashtags
    - Avg. ~6 tags per tweet

- Pivot to Users x Tags

- 60k Users x 9k Unique Tags
    - Filter at least 9 interactions (tags / uses)
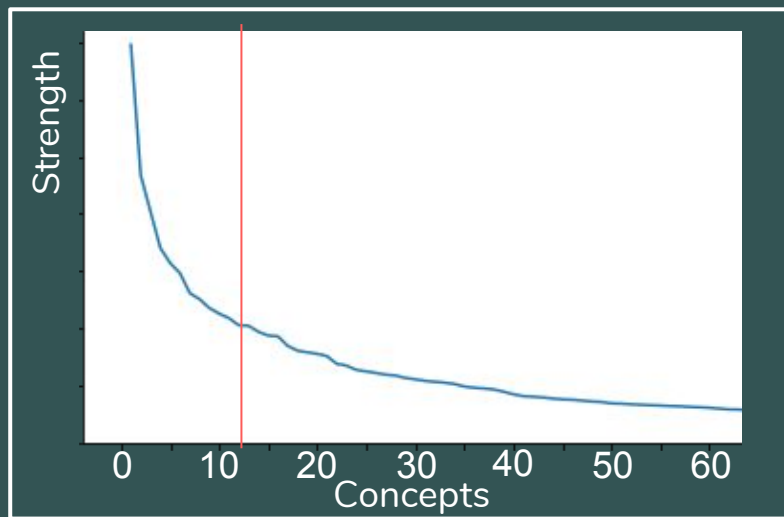   - Avg User: ~10 unique tags, ~6 tweets

Recommender Models

- Popularity (baseline):
  - Recommend top 10 hashtags not yet tweeted

- Collaborative:
  - Find 5 most similar users (SVD - cosine)
  - Recommend most tweeted by similar users not yet tweeted

Train - Test Methodology

- Train (80% of users)

- Test (20% of users)
  - Metric: Recall@N
  - For each test user:
    Create profile with one hashtag held out
    Generate recommendations
    Recall = % of time holdout appears in Top N

Detail: Collaborative Model
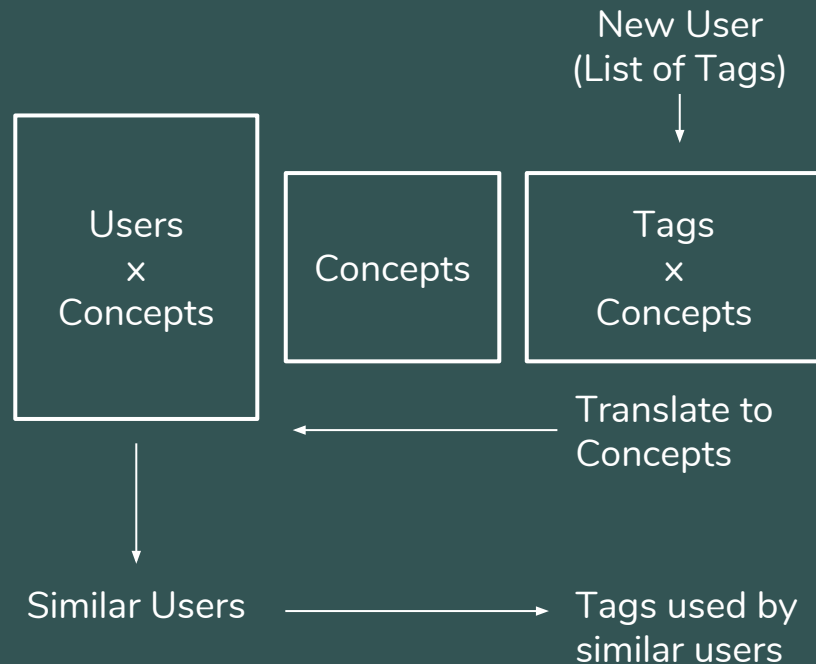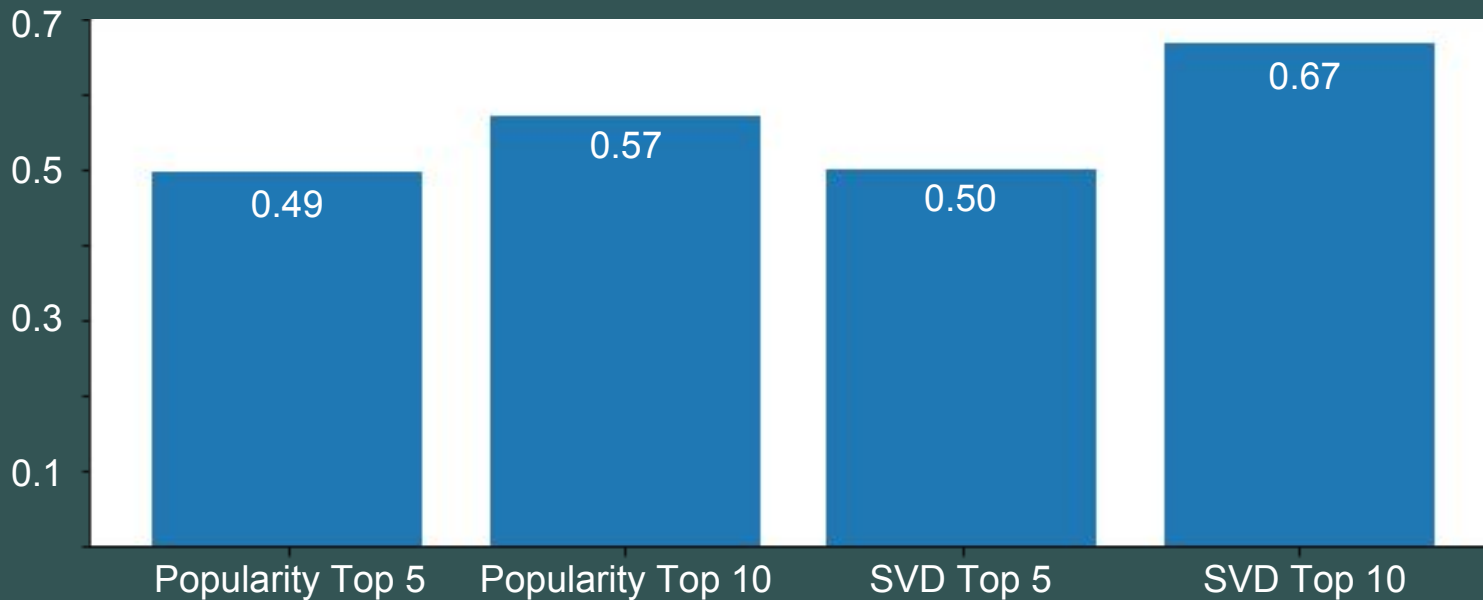
SVD: 13 Concepts

<-Generalization   Precision->

New User
(List of Tags)

Users × Concepts	Concepts	Tags × Concepts

Translate to Concepts

Similar Users → Tags used by similar users

Train   Test

Test Results

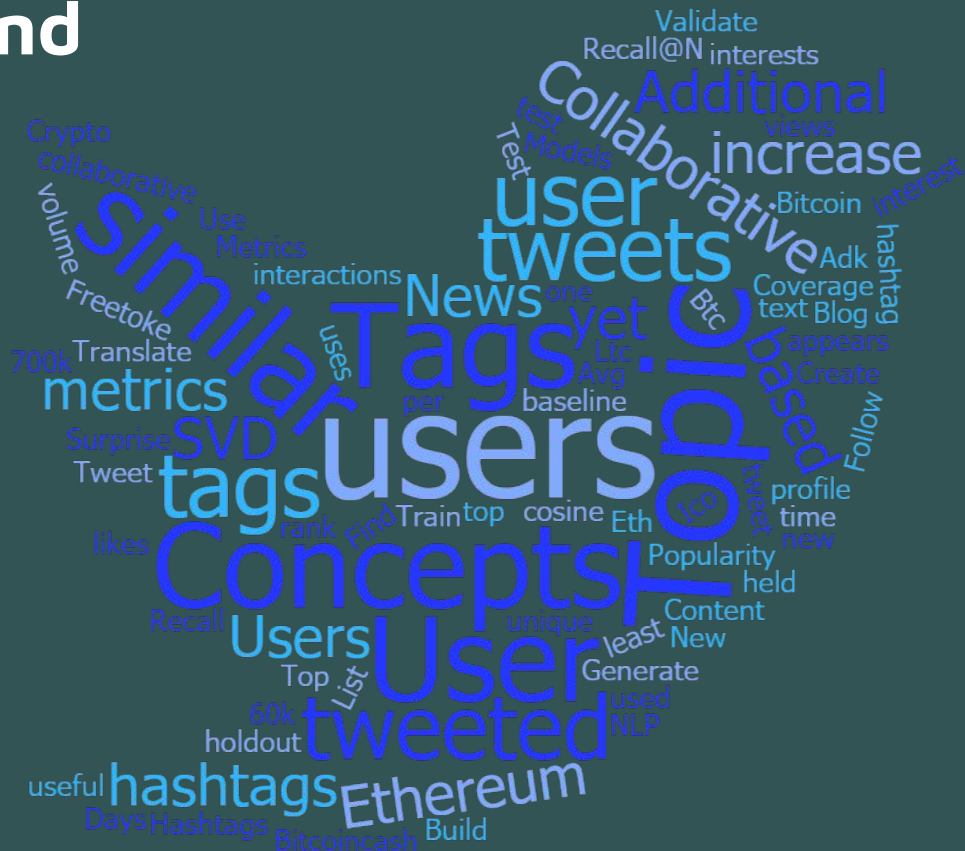Manual review also looked reasonable but hard to differentiate tags

Next Steps

Additional Models

- Other collaborative

- Content based

- Predict numeric ratings (regression)

Additional Metrics

- $ / Response rate

- Coverage / Surprise

The End

Processing... NLP on Text?

NMF topics: Tweets - Hashtags

Topic 1:
bestboyband: 8.258
Iheartawards: 8.180
Bts: 7.943
Thosfansbts4m: 0.261

Topic 2:
Bitcoin: 12.999
News: 0.288
Ethereum: 0.184
Bitcoincash: 0.169

Topic 3:
Ethereum: 4.109
pumpkinpay: 3.817
cryptobountyprogram: 3.805
Freetoke: 3.791

Topic 4:
Btc: 8.895
Eth: 0.869
Ltc: 0.376
Telegramico: 0.304

Topic 5:
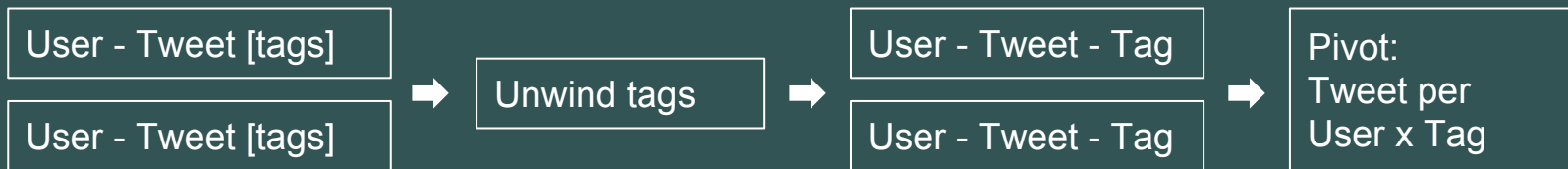Crypto: 7.404
News: 0.444
Aidoskuneen: 0.328
Adk: 0.328

Topic 6:
Blockchain: 7.183
Ico: 0.400
Fintech: 0.369
Linonetwork: 0.251

Processing... User-Hashtag Matrix

User - Tweet [tags]

User - Tweet [tags]

➡️

Unwind tags

➡️

User - Tweet - Tag

User - Tweet - Tag

➡️

Pivot: Tweet per User x Tag

Columns: Hashtags (h1, h2, ..., hn)

Rows: Users (u1, u2, ..., um)

Values: count uses of hashtag h by user