

Subreddit Recommender

Vince Kumar

Data Scientist

[Github.com/VinceKumar](https://github.com/VinceKumar)

<https://www.linkedin.com/in/vince-kumar>

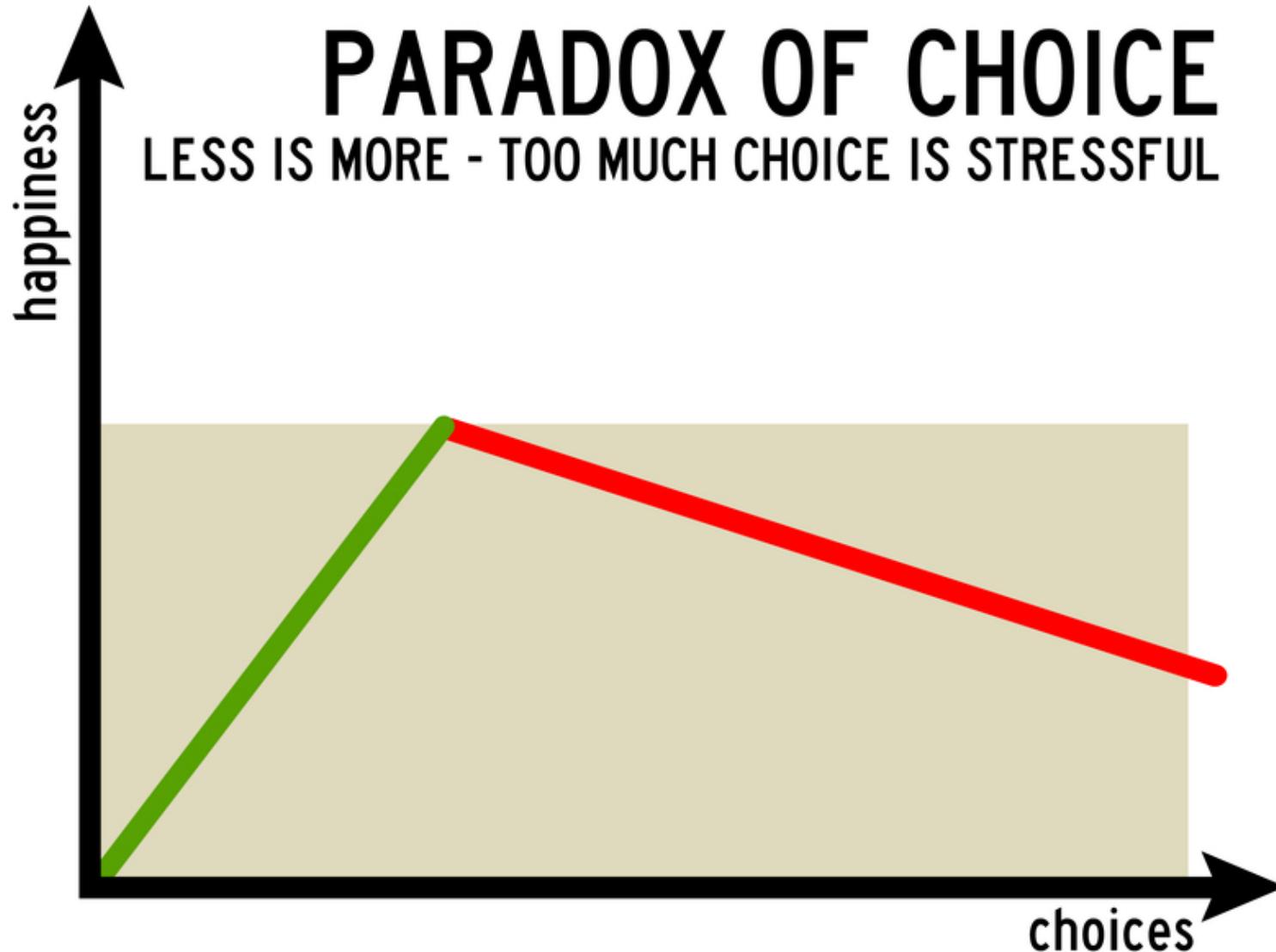


What is Reddit?

- 4th Most visited website in the US*
- Each subreddit is an online community
- Reddit currently has no recommender

r/news
r/books
r/movies
r/politics
r/science
r/Music
r/technology

Too few choices are not good enough, and
too many choices is overwhelming

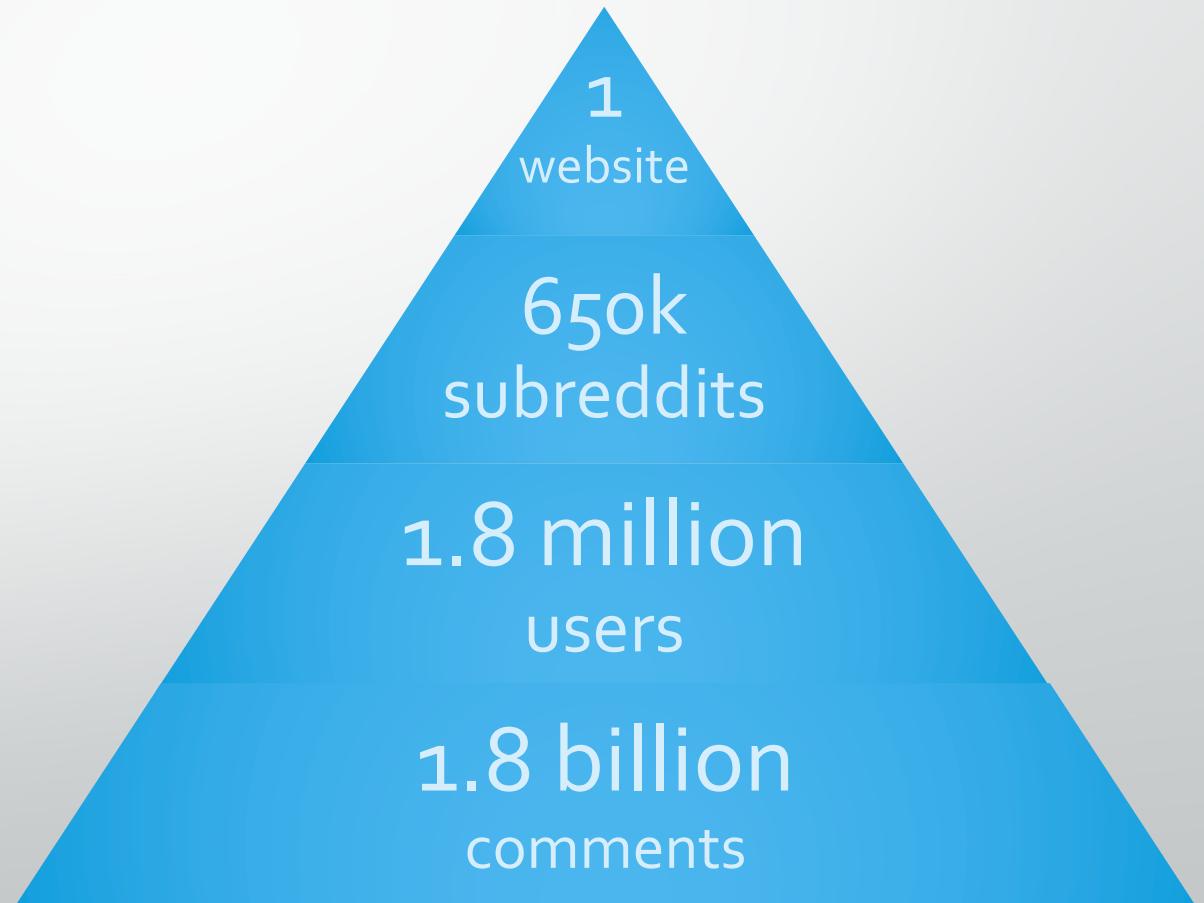


Dataset

1 TB Dataset

Found on r/datasets

8 years of data



Approach

Collaborative Filtering (User Similarity)

	Machine Learning	aww	programming	Deep Learning	puppies	compsci	technews
Gavin	Yes			Yes			
Richard	Yes		Yes	Yes		Yes	Yes
Nelson	Yes						Yes
Gilfoyle	Yes		Yes				Yes

Yes = comment activity
in that subreddit



	Most Similar	2 nd Most Similar	
Gavin	Richard	Nelson	Gilfoyle



	Top Recommendation	2 nd Recommendation	3 rd Recommendation
Gavin	technews	programming	compsci

Approach

Collaborative Filtering (User Similarity)

	Machine Learning	aww	programming	Deep Learning	puppies	compsci	technews
Gavin	Yes			Yes			
Richard	Yes		Yes	Yes		Yes	Yes
Nelson	Yes						Yes
Gilfoyle	Yes		Yes				Yes

Yes = comment activity
in that subreddit



3 votes

	Most Similar	2 nd Most Similar	3 rd Most Similar
Gavin	Richard	Nelson	Gilfoyle



	Top Recommendation	2 nd Recommendation	3 rd Recommendation
Gavin	technews	programming	compsci

Approach

Collaborative Filtering (User Similarity)

	Machine Learning	aww	programming	Deep Learning	puppies	compsci	technews
Gavin	Yes			Yes			
Richard	Yes		Yes	Yes	Yes	Yes	Yes
Nelson	Yes						Yes
Gilfoyle	Yes		Yes				Yes

Yes = comment activity
in that subreddit

2 votes

3 votes

	Most Similar	2 nd Most Similar	3 rd Most Similar
Gavin	Richard	Nelson	Gilfoyle



	Top Recommendation	2 nd Recommendation	3 rd Recommendation
Gavin	technews	programming	compsci

Approach

Collaborative Filtering (User Similarity)

Yes = comment activity
in that subreddit

	Machine Learning	aww	programming	Deep Learning	puppies	compsci	technews
Gavin	Yes			Yes			
Richard	Yes		Yes	Yes		Yes	Yes
Nelson	Yes						Yes
Gilfoyle	Yes		Yes				Yes

2 votes



1 vote

3 votes

	Most Similar	2 nd Most Similar	3 rd Most Similar
Gavin	Richard	Nelson	Gilfoyle



	Top Recommendation	2 nd Recommendation	3 rd Recommendation
Gavin	technews	programming	compsci

Evaluation

- Difficult to evaluate an unsupervised model



What Next?

**Implement Content-Based,
Collaborative Recommendation
(Hybrid)**

Contact Me

- Email: VinceKumar8@Gmail.com
- Github: [VinceKumar](https://github.com/VinceKumar)
- Linkedin: [Linkedin.com/in/vince-kumar](https://www.linkedin.com/in/vince-kumar)

