Quora

LinkedIn

Lessons Learned from Building real-life Recsys

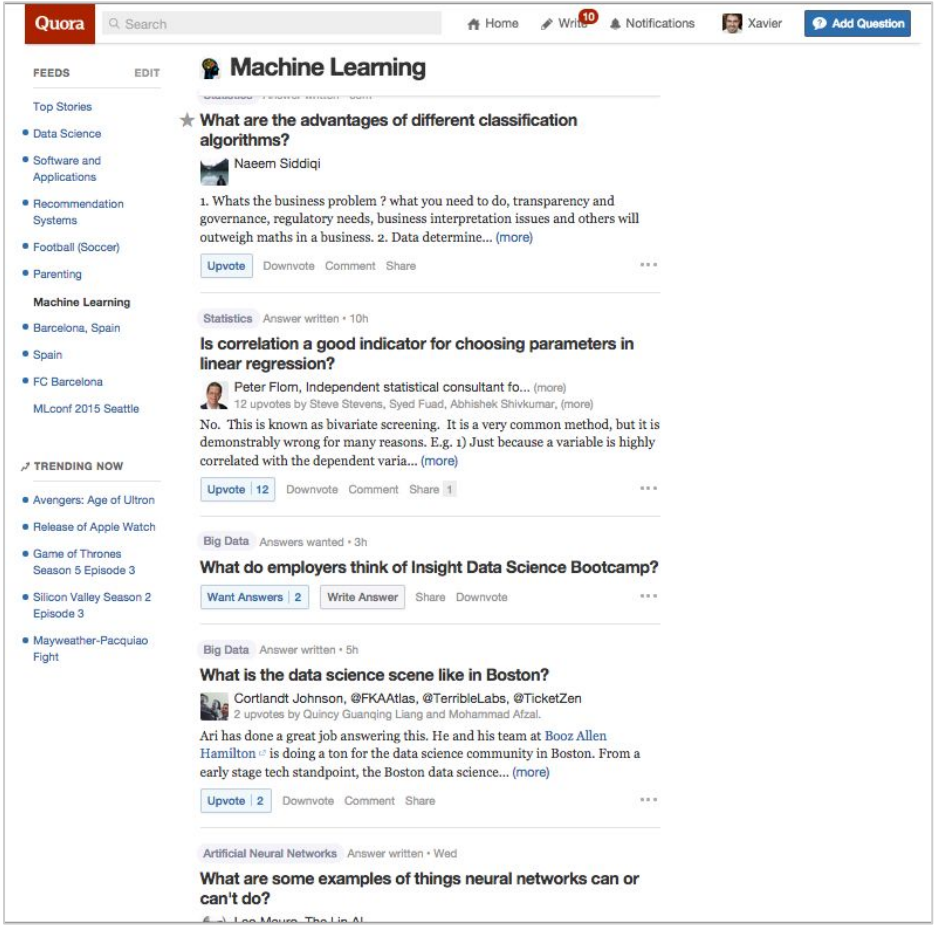
Xavier Amatriain (Quora)
Deepak Agarwal (LinkedIn)

4. Recommendations @Quora

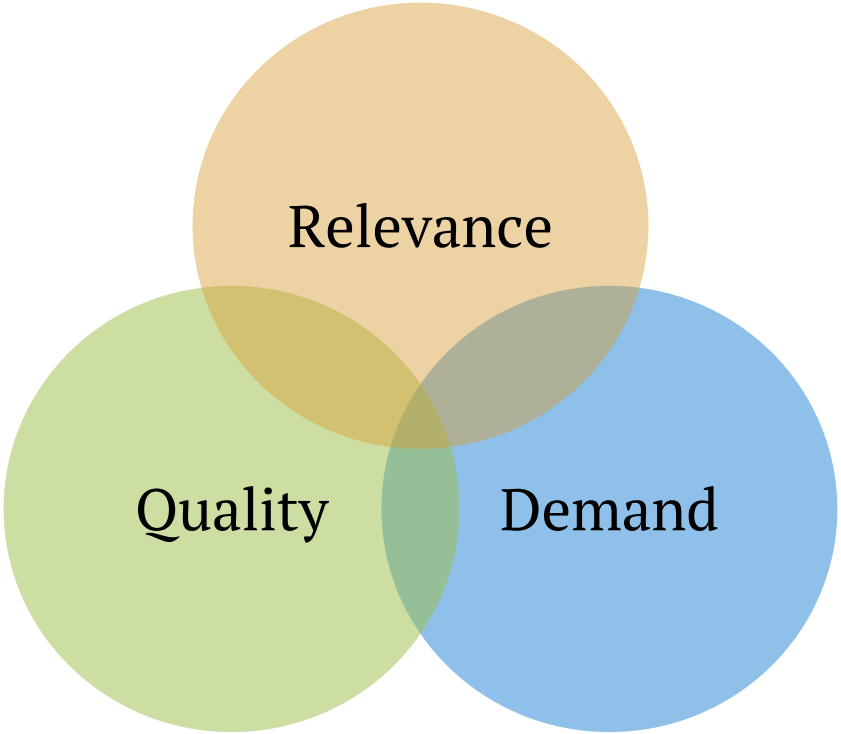
Our Mission

“To share and grow the world’s knowledge”

- Millions of questions & answers
- Millions of users
- Thousands of topics
- …

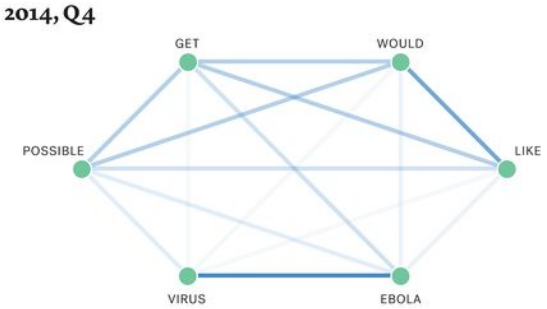


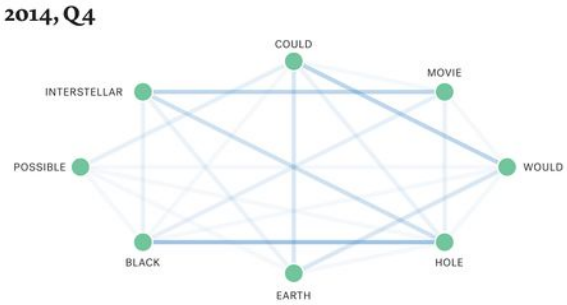
What we care about



Data @ Quora

Lots of high-quality textual information

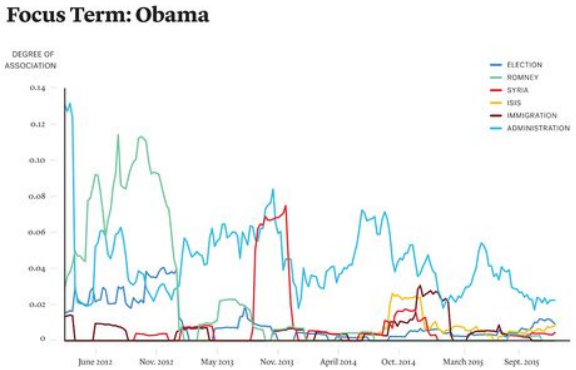




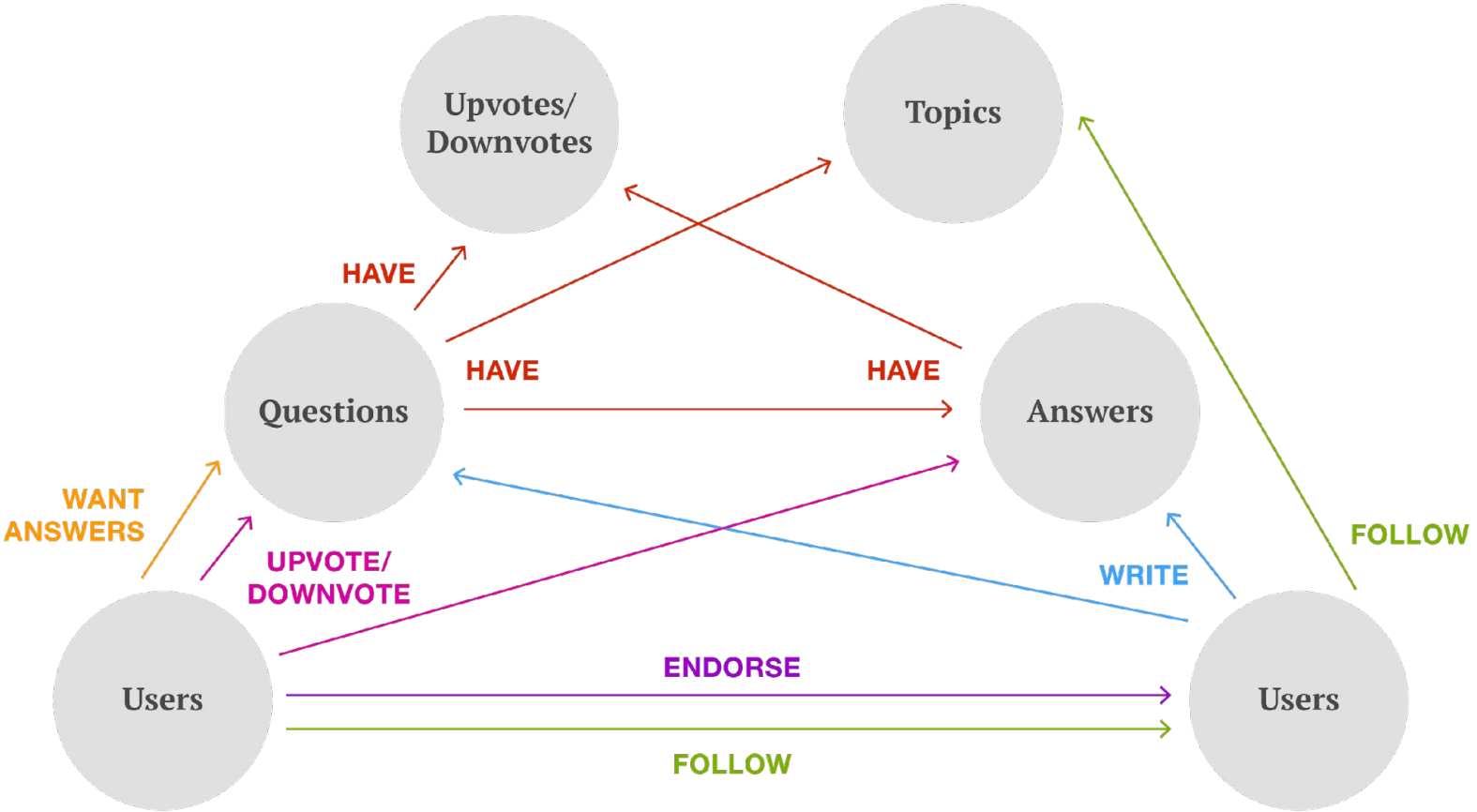
DATA @ QUORA

Mapping the Discussion on Quora Over Time through Question Text

WENWEN TAO (陶雯雯)



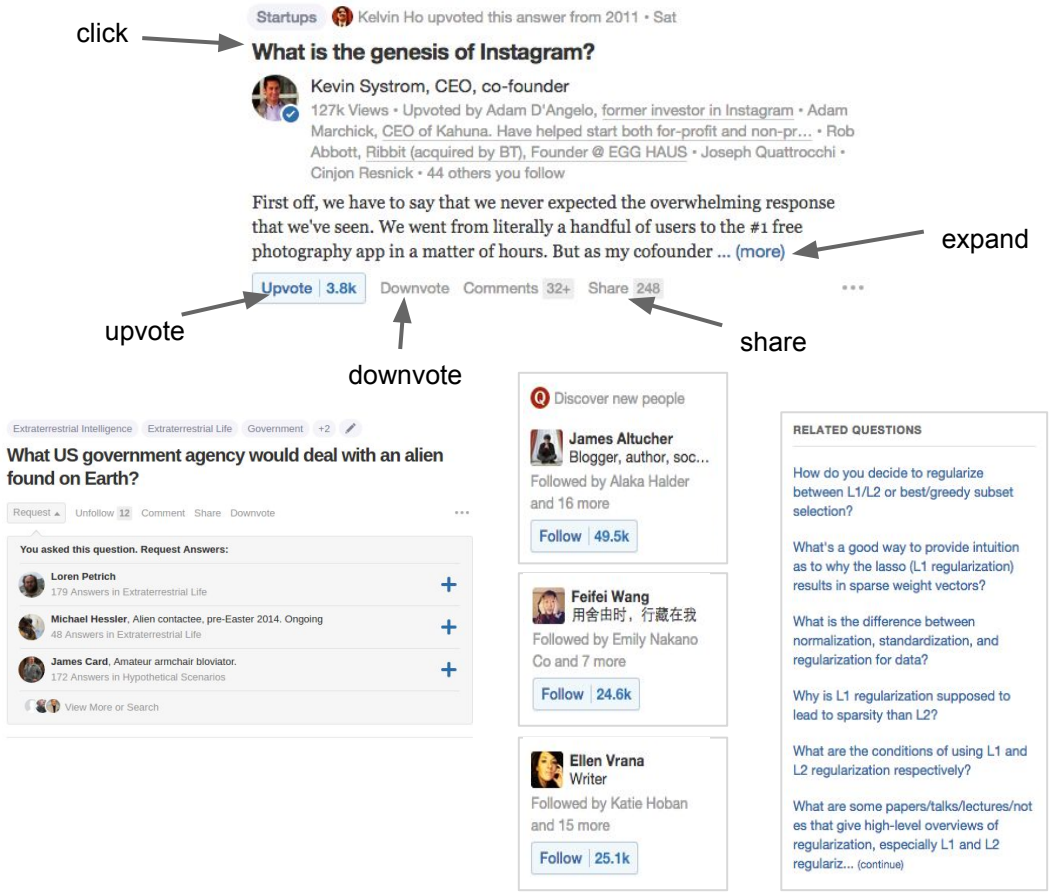
Lots of data relations



Recommendations @Quora

Recommendations at Quora

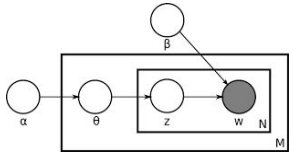
- Homepage feed ranking
- Email digest
- Answer ranking
- Topic recommendation
- User recommendation
- Trending Topics
- Automated Topic Labelling
- Related Question
- ...



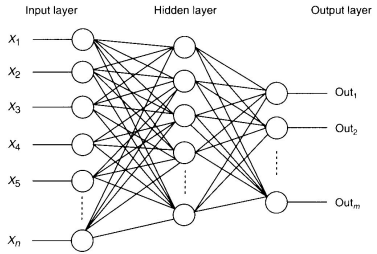
Models

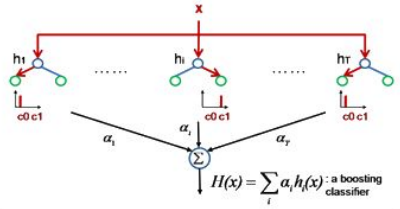
Quora

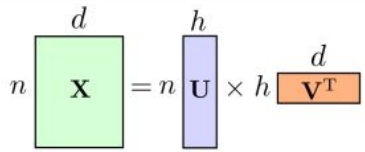
- Deep Neural Networks
- Logistic Regression
- Elastic Nets
- Gradient Boosted Decision Trees
- Random Forests
- LambdaMART
- Matrix Factorization
- LDA
- ...
-

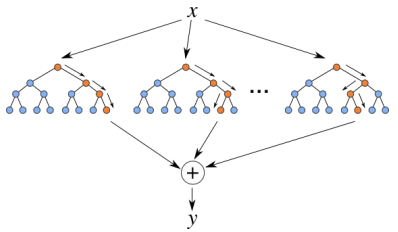


$$P = \frac{e^{a+bX}}{1+e^{a+bX}}$$









$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}}(\|y - X\beta\|^2 + \lambda_2\|\beta\|^2 + \lambda_1\|\beta\|_1).$$

5. Lessons learned (Part II)

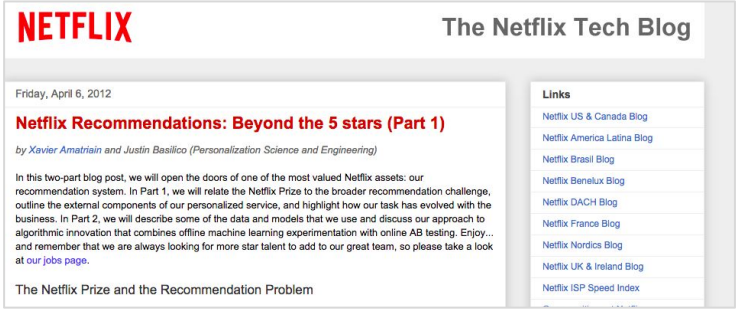
1. Implicit signals beat explicit ones (almost always)

Quora

Implicit vs. Explicit

Quora

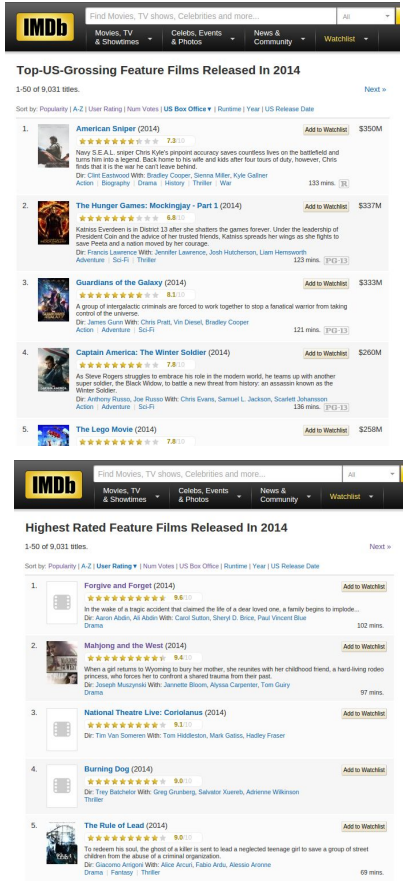
- Many have acknowledged that implicit feedback is more useful
- Is implicit feedback really always more useful?
- If so, why?



Implicit vs. Explicit

Quora

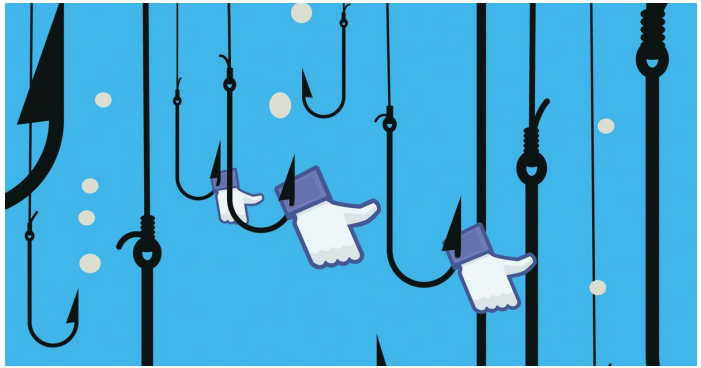
- Implicit data is (usually):
  - More dense, and available for all users
  - Better representative of user behavior vs. user reflection
  - More related to final objective function
  - Better correlated with AB test results
- E.g. Rating vs watching

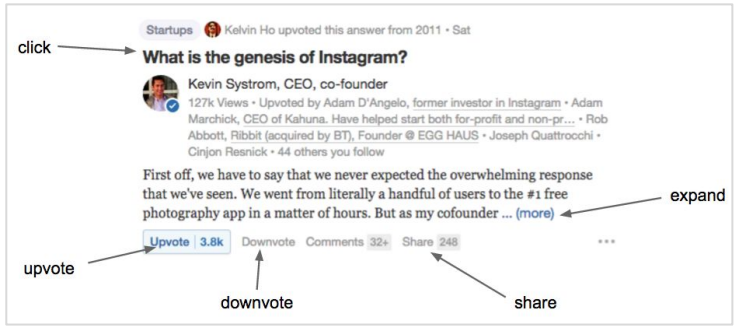


Implicit vs. Explicit

Quora

- However
  - It is not always the case that direct implicit feedback correlates well with long-term retention
  - E.g. clickbait
- Solution:
  - Combine different forms of implicit + explicit to better represent long-term goal





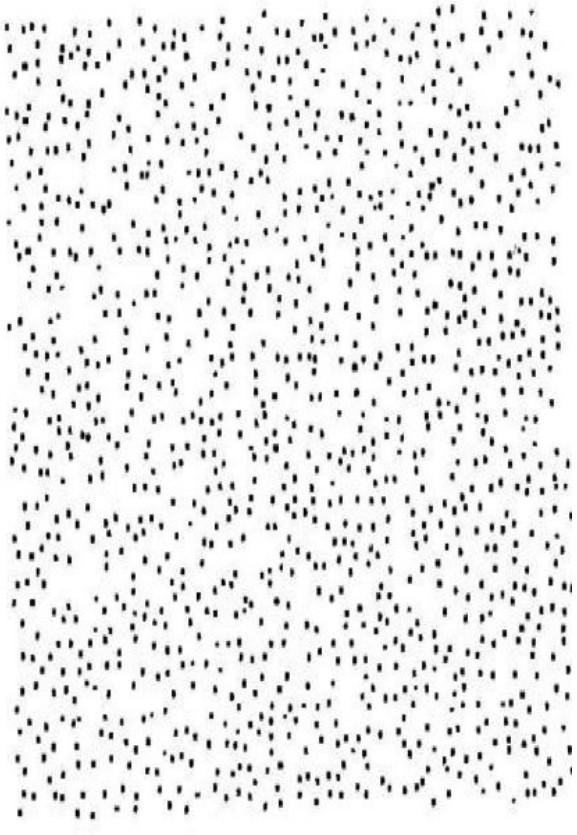
2. Be thoughtful about your Training Data

Quora

Defining training/testing data

Quora

- Training a simple binary classifier for good/bad answer
  - Defining positive and negative labels -> Non-trivial task
  - *Is this a positive or a negative?*
    - funny uninformative answer with many upvotes
    - short uninformative answer by a well-known expert in the field
    - very long informative answer that nobody reads/upvotes
    - informative answer with grammar/spelling mistakes
    - ...



3. Your Model will learn what you teach it to learn

Quora

Training a model

Quora

- Model will learn according to:
  - Training data (e.g. implicit and explicit)
  - Target function (e.g. probability of user reading an answer)
  - Metric (e.g. precision vs. recall)
- Example 1 (made up):
  - *Optimize probability of a user going to the cinema to watch a movie and rate it “highly” by using purchase history and previous ratings. Use NDCG of the ranking as final metric using only movies rated 4 or higher as positives.*

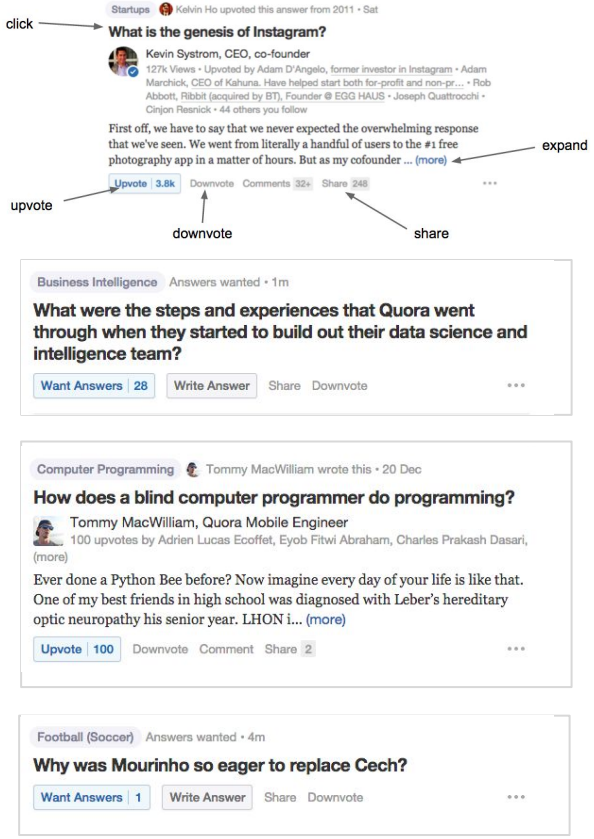
Example 2 - Quora's feed

Quora

- Training data = implicit + explicit
- Target function: Value of showing a story to a user ~ weighted sum of actions:

$$v = \sum_a v_a \, 1\{y_a = 1\}$$

  - predict probabilities for each action, then compute expected value: $v\_pred = E[\, V \mid x \,] = \sum_a v_a \, p(a \mid x)$
- Metric: any ranking metric



4. Explanations might matter more than the prediction

Quora

Explanation/Support for Recommendations

Quora

Sarah Smith Richard Henry and 3 more upvoted this • 7h

How can I complain about my roommate who is cheating on his Google phone interviews?

Ben Garrison, Software Engineer at Google
304.3k Views • Upvoted by Jeremy Miles, Quantitative analyst at Google, Mayeesha Tahsin, Sarah Smith, and 3 others you follow

First off, I really appreciate your trying to make sure the right thing happens. I think that's great. Cheating sucks. However, the answer is "don't worry about it". Phone screens here at Google ar... (more)

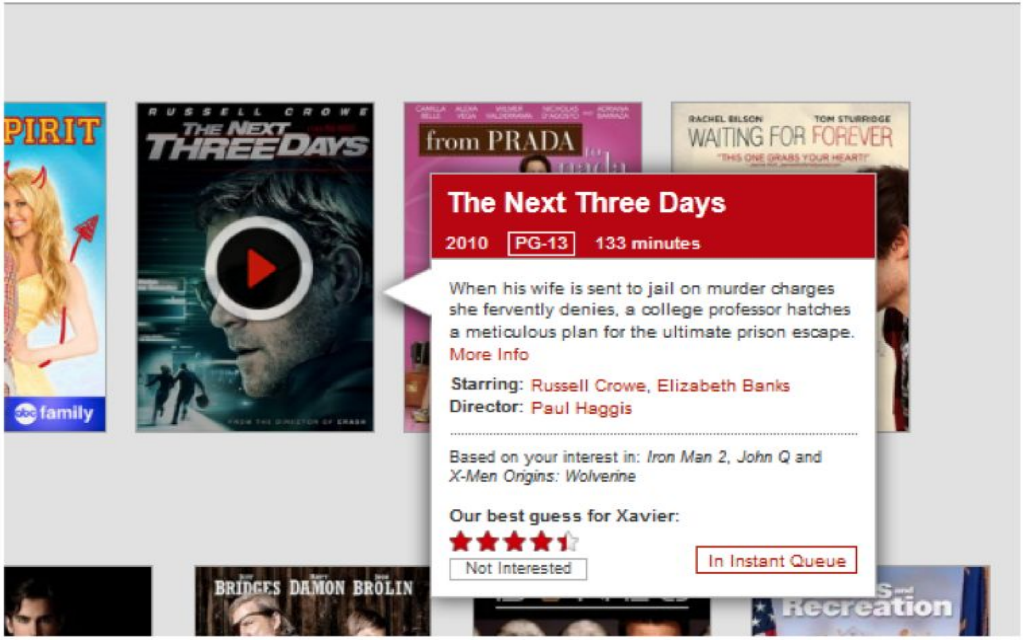
Upvote 968 Downvote Comments 23+ Share

Discover new topics

Last.fm
Last.fm builds detail...
Followed by Neal Lathia and 8 more
Follow 21.9k

Quantitative Finance
Quantitative finance ...
Followed by Katie Hoban and 22 more
Follow 74.1k

California Stat
California State
Followed by Rachelle Baratto
Follow 2.4k



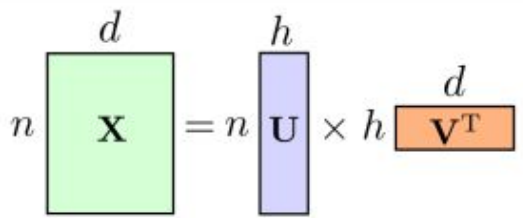
5. If You Have to Pick one single approach, Matrix factorization is your best bet

Quora

Matrix Factorization

Quora

- MF can be interpreted as
  - Unsupervised:
    - Dimensionality Reduction a la PCA
    - Clustering (e.g. NMF)
  - Supervised:
    - Labeled targets ~ regression



- Very useful variations of MF
  - BPR, ALS, SVD++
  - Tensor Factorization, Factorization Machines
- However...

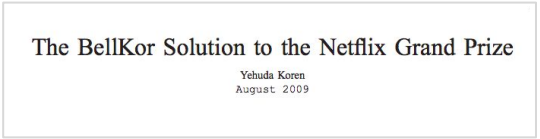
6. Everything is an ensemble

Quora

Ensembles

Quora

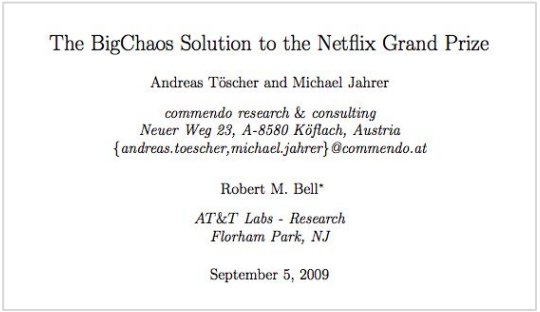
- Netflix Prize was won by an ensemble
  - Initially Bellkor was using GDBTs
  - BigChaos introduced ANN-based ensemble
- Most practical applications of ML run an ensemble
  - Why wouldn't you?
  - At least as good as the best of your methods
  - Can add completely different approaches (e.g. CF and content-based)
  - You can use many different models at the ensemble layer: LR, GDBTs, RFs, ANNs…



The BellKor Solution to the Netflix Grand Prize

Yehuda Koren

August 2009



The BigChaos Solution to the Netflix Grand Prize

Andreas Töscher and Michael Jahrer

commendo research & consulting
Neuer Weg 23, A-8580 Köflach, Austria
{andreas.toescher,michael.jahrer}@commendo.at

Robert M. Bell*

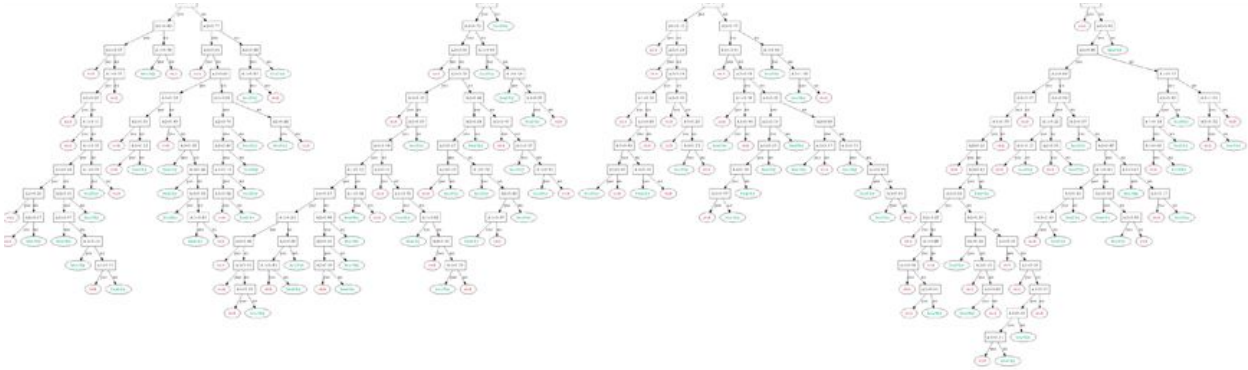
AT&T Labs - Research
Florham Park, NJ

September 5, 2009

Ensembles & Feature Engineering

Quora

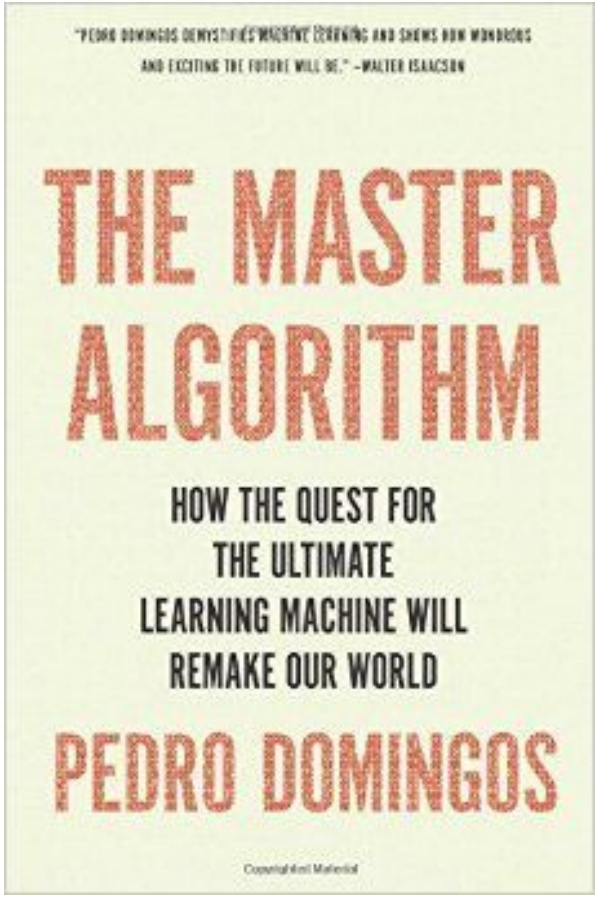
- Ensembles are the way to turn any model into a feature!
- E.g. Don’t know if the way to go is to use Factorization Machines, Tensor Factorization, or RNNs?
  - Treat each model as a “feature”
  - Feed them into an ensemble



The Master Algorithm?

Quora

It definitely is an ensemble!



7. Building Recommender Systems is also about Feature Engineering

Quora

Feature Engineering

Quora

Need for feature engineering

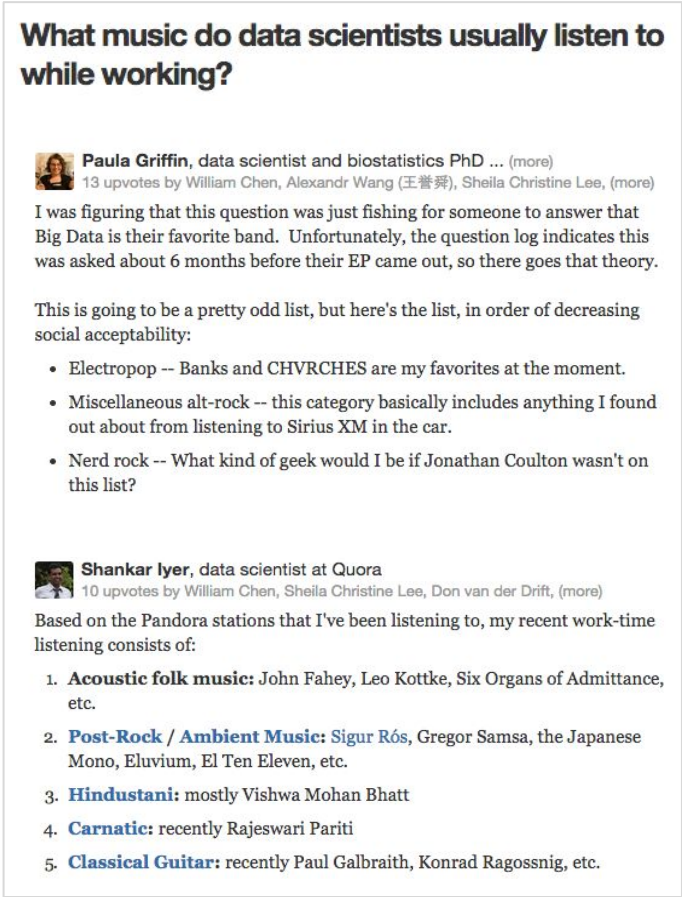
In many cases an understanding of the domain will lead to optimal results.

Feature Engineering Example - Quora Answer Ranking

Quora

What is a good Quora answer?

- truthful
- reusable
- provides explanation
- well formatted
- ...



What music do data scientists usually listen to while working?

Paula Griffin, data scientist and biostatistics PhD ... (more)
13 upvotes by William Chen, Alexandr Wang (王誉舜), Sheila Christine Lee, (more)

I was figuring that this question was just fishing for someone to answer that Big Data is their favorite band. Unfortunately, the question log indicates this was asked about 6 months before their EP came out, so there goes that theory.

This is going to be a pretty odd list, but here's the list, in order of decreasing social acceptability:

- Electropop -- Banks and CHVRCHES are my favorites at the moment.
- Miscellaneous alt-rock -- this category basically includes anything I found out about from listening to Sirius XM in the car.
- Nerd rock -- What kind of geek would I be if Jonathan Coulton wasn't on this list?

Shankar Iyer, data scientist at Quora
10 upvotes by William Chen, Sheila Christine Lee, Don van der Drift, (more)

Based on the Pandora stations that I've been listening to, my recent work-time listening consists of:

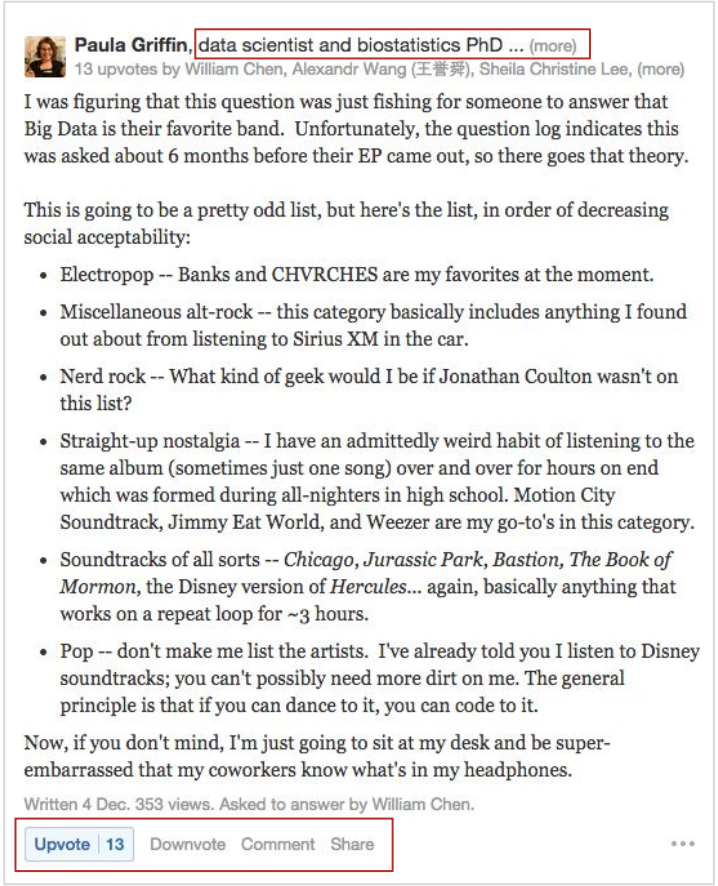
1. **Acoustic folk music:** John Fahey, Leo Kottke, Six Organs of Admittance, etc.
2. **Post-Rock / Ambient Music:** Sigur Rós, Gregor Samsa, the Japanese Mono, Eluvium, El Ten Eleven, etc.
3. **Hindustani:** mostly Vishwa Mohan Bhatt
4. **Carnatic:** recently Rajeswari Pariti
5. **Classical Guitar:** recently Paul Galbraith, Konrad Ragossnig, etc.

Feature Engineering Example - Quora Answer Ranking

Quora

How are those dimensions translated into features?

- Features that relate to the answer quality itself
- Interaction features (upvotes/downvotes, clicks, comments...)
- User features (e.g. expertise in topic)



Paula Griffin, data scientist and biostatistics PhD ... (more)
13 upvotes by William Chen, Alexandr Wang (王誉舜), Sheila Christine Lee, (more)

I was figuring that this question was just fishing for someone to answer that Big Data is their favorite band. Unfortunately, the question log indicates this was asked about 6 months before their EP came out, so there goes that theory.

This is going to be a pretty odd list, but here's the list, in order of decreasing social acceptability:

- Electropop -- Banks and CHVRCHES are my favorites at the moment.
- Miscellaneous alt-rock -- this category basically includes anything I found out about from listening to Sirius XM in the car.
- Nerd rock -- What kind of geek would I be if Jonathan Coulton wasn't on this list?
- Straight-up nostalgia -- I have an admittedly weird habit of listening to the same album (sometimes just one song) over and over for hours on end which was formed during all-nighters in high school. Motion City Soundtrack, Jimmy Eat World, and Weezer are my go-to's in this category.
- Soundtracks of all sorts -- *Chicago*, *Jurassic Park*, *Bastion*, *The Book of Mormon*, the Disney version of *Hercules*... again, basically anything that works on a repeat loop for ~3 hours.
- Pop -- don't make me list the artists. I've already told you I listen to Disney soundtracks; you can't possibly need more dirt on me. The general principle is that if you can dance to it, you can code to it.

Now, if you don't mind, I'm just going to sit at my desk and be super-embarrassed that my coworkers know what's in my headphones.

Written 4 Dec. 353 views. Asked to answer by William Chen.

Upvote 13 Downvote Comment Share ...

Feature Engineering

Quora

- Properties of a well-behaved ML feature:
  - Reusable
  - Transformable
  - Interpretable
  - Reliable

Deep Learning

NIPS'2015 Tutorial

Geoff Hinton, Yoshua Bengio & Yann LeCun

CIFAR
CANADIAN INSTITUTE for ADVANCED RESEARCH

Deep Learning: Automating Feature Discovery

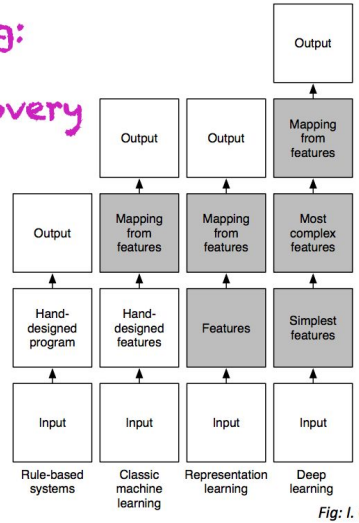


Fig: I. Goodfellow

« Smerity.com

In deep learning, architecture engineering is the new feature engineering

June 11, 2016

Two of the most important aspects of machine learning models are feature extraction and feature engineering. Those features are what supply relevant information to the machine learning models.

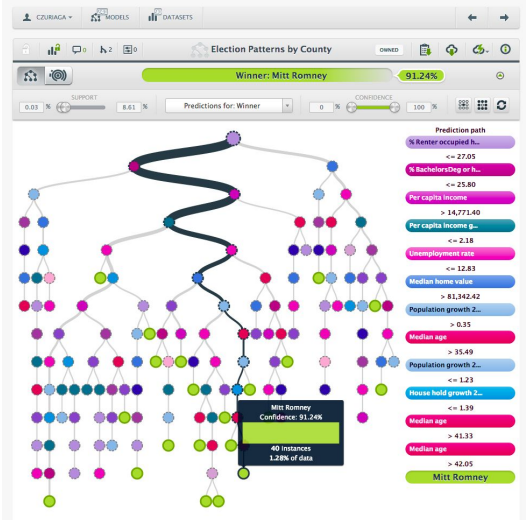
8. Why you should care about answering questions (about your recsys)

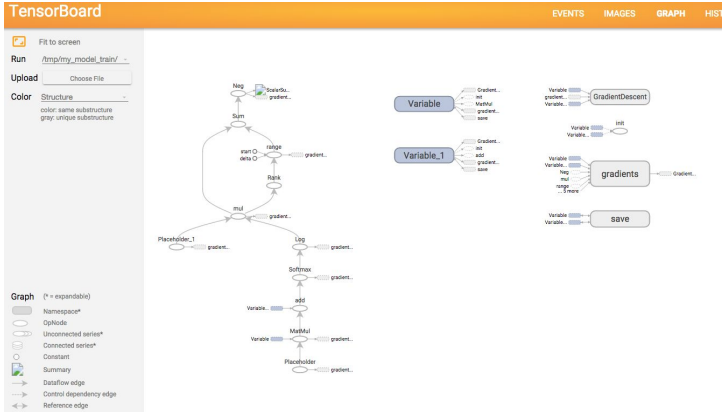
Quora

Model debuggability

Quora

- Value of a model = value it brings to the product
- Product owners/stakeholders have expectations on the product
- It is important to answer questions to why did something fail
- Bridge gap between product design and ML algos
- Model debuggability is so important it can determine:
  - Particular model to use
  - Features to rely on
  - Implementation of tools

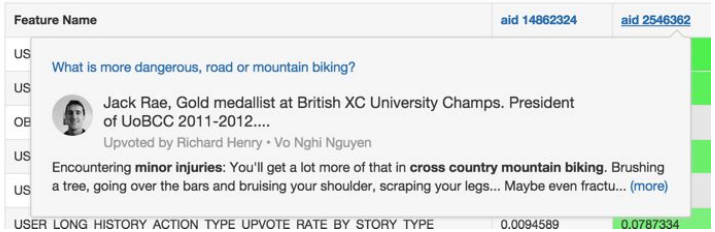


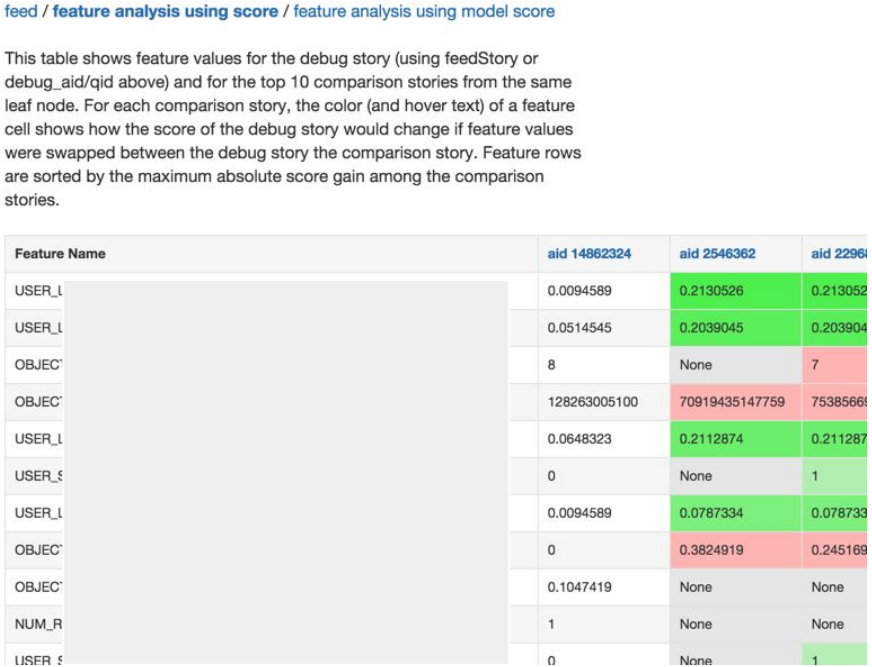


Model debuggability

Quora

- E.g. Why am I seeing or not seeing this on my homepage feed?





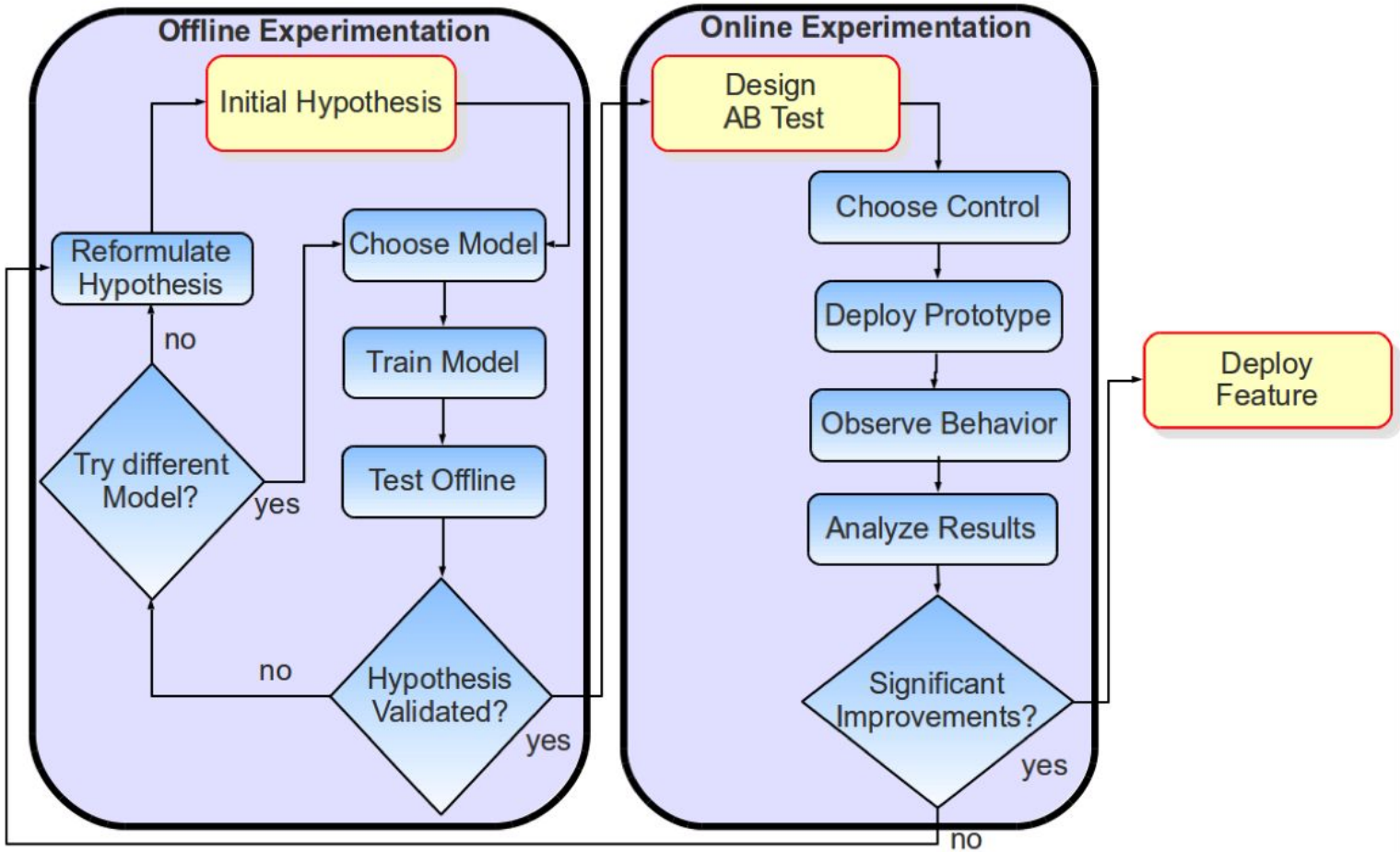
9. Data and Models are great. You know what's even better?

The right evaluation approach!

Quora

Offline/Online testing process

Quora



Executing A/B tests

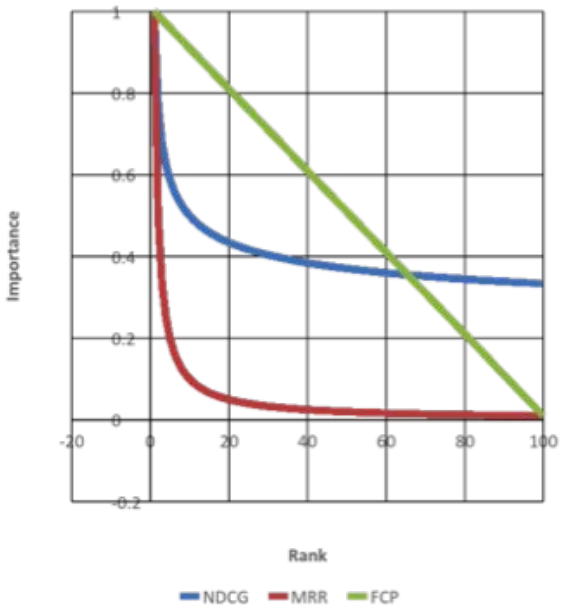
Quora

- **Measure *differences* in metrics across *statistically identical* populations that each experience a different algorithm.**
- Decisions on the product always data-driven
- Overall Evaluation Criteria (OEC) = member retention
  - Use long-term metrics whenever possible
  - Short-term metrics can be informative and allow faster decisions
    - But, not always aligned with OEC

Offline testing

Quora

- Measure model performance, using (IR) metrics
- Offline performance = indication to make decisions on follow-up A/B tests
- A critical (and mostly unsolved) issue is how offline metrics correlate with A/B test results.



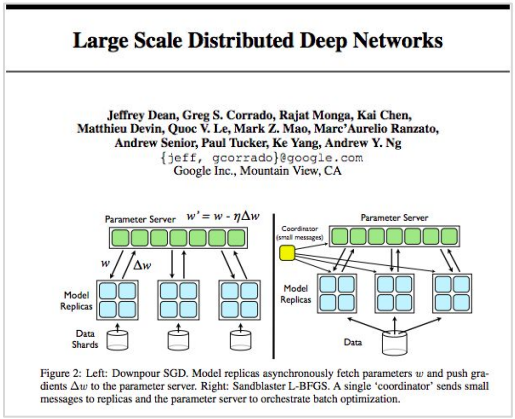
10. You don't need to distribute your Recsys

Quora

Distributing Recommender Systems

Quora

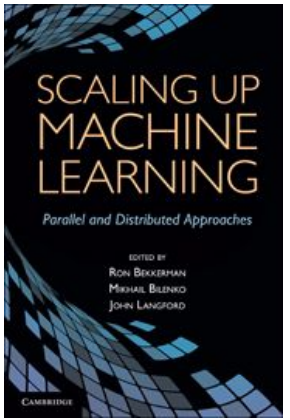
- Most of what people do in practice can fit into a multi-core machine
  - As long as you use:
    - Smart data sampling
    - Offline schemes
    - Efficient parallel code
- (... but not Deep ANNs)



Large Scale Distributed Deep Networks

Jeffrey Dean, Greg S. Corrado, Rajat Monga, Kai Chen,
Matthieu Devin, Quoc V. Le, Mark Z. Mao, Marc'Aurelio Ranzato,
Andrew Senior, Paul Tucker, Ke Yang, Andrew Y. Ng
{jeff, gcorrado}@google.com
Google Inc., Mountain View, CA

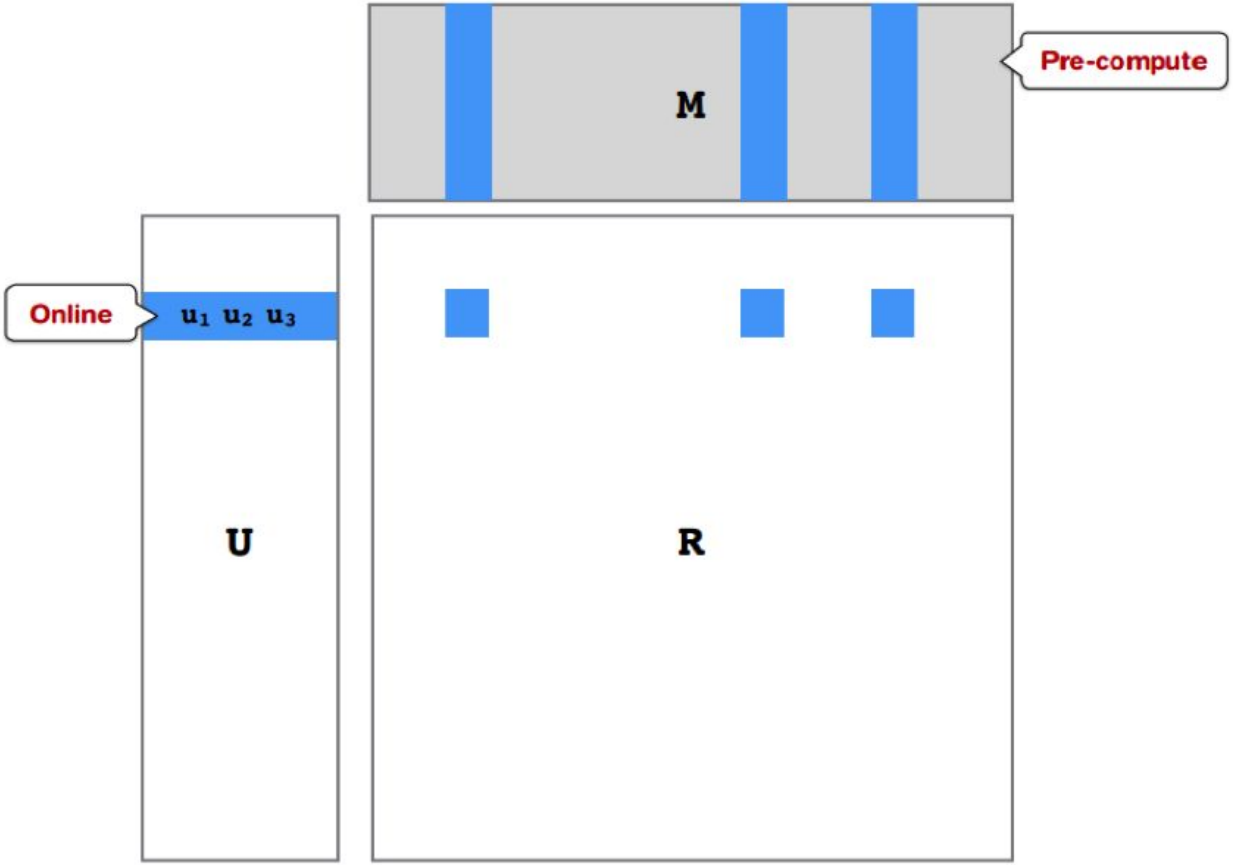
Figure 2: Left: Downpour SGD. Model replicas asynchronously fetch parameters w and push gradients Δw to the parameter server. Right: Sandblaster L-BFGS. A single 'coordinator' sends small messages to replicas and the parameter server to orchestrate batch optimization.



- Do you care about costs? How about latencies or system complexity/debuggability?

Matrix Factorization Example

Quora



Conclusions

- Recommender Systems are about much more than just predicting a rating
- Designing a “real-life” recsys means paying attention to issues such as:
  - Feature engineering
  - Training dataset
  - Metrics
  - Experimentation and AB Testing
  - System scalability
  - ...
- Lots of room for improvement & research

Questions?

Xavier Amatriain (Quora)
xavier@amatriain.net

Deepak Agarwal (LinkedIn)
dagarwal@linkedin.com

Quora

LinkedIn