

Designing Evaluation

Match the Evaluation to the Question, Problem, or Challenge

- What's the right metric?
- How should I sample the data?
- How do I know the right algorithm to use?
- The answer is always the same!!!

An Example: News Recommendation

- Let's consider Google News
- What are the questions we might have about its design and performance?
 - are we recommending useful articles?
 - are there ways we should improve the recommender?
 - Diversity, Serendipity, Familiarity, Level of Perceived Personalization, Balkanization?

Let's talk through these ...

- Useful recommendations
- Potential improvements

Key lesson ...

- It depends ...
- Evaluation requires its own design
- Up next ... two more examples!

Designing Evaluation