Building a Dynamic, Contextual Recommendation System

Ryan Dew

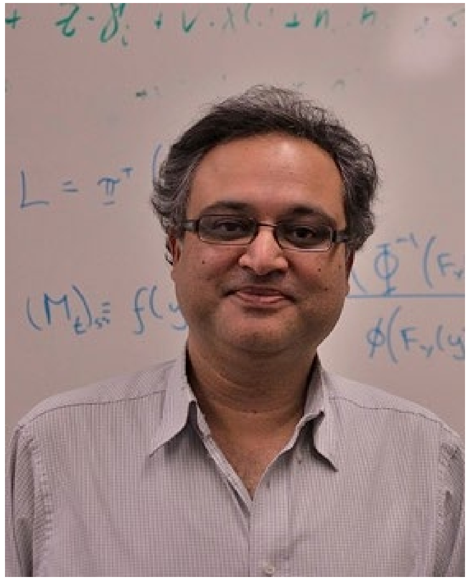
The Wharton School, University of Pennsylvania

WCAI Annual Conference 2019

The Team

Ryan Dew
Assistant Professor of Marketing
The Wharton School,
University of Pennsylvania

Yegor Tkachenko
Ph.D. Student
Columbia Business School,
Columbia University



Asim Ansari
William T. Dillard Professor of Marketing
Columbia Business School,
Columbia University

The Problem: Researcher's Perspective

The goal: For any player, understand his/her preferences over all possible games

User	Date	Game
1	2016-06-10	Game 1
1	2016-06-11	Game 1
1	2016-06-13	Game 1
1	2016-06-13	Game 2
	⋮	
3	2016-06-10	Game 5
	⋮	

User 1: Game 1 > 2 > 5 > 3 > …
User 3: Game 5 > 7 > 2 > 3 > …

The Problem: Researcher's Perspective

The goal: For any player, understand his/her preferences over all possible games

Key challenges:

- Providing EA *useful outputs*

User	Date	Game
1	2016-06-10	Game 1
1	2016-06-11	Game 1
1	2016-06-13	Game 1
1	2016-06-13	Game 2
	⋮	
3	2016-06-10	Game 5
	⋮	

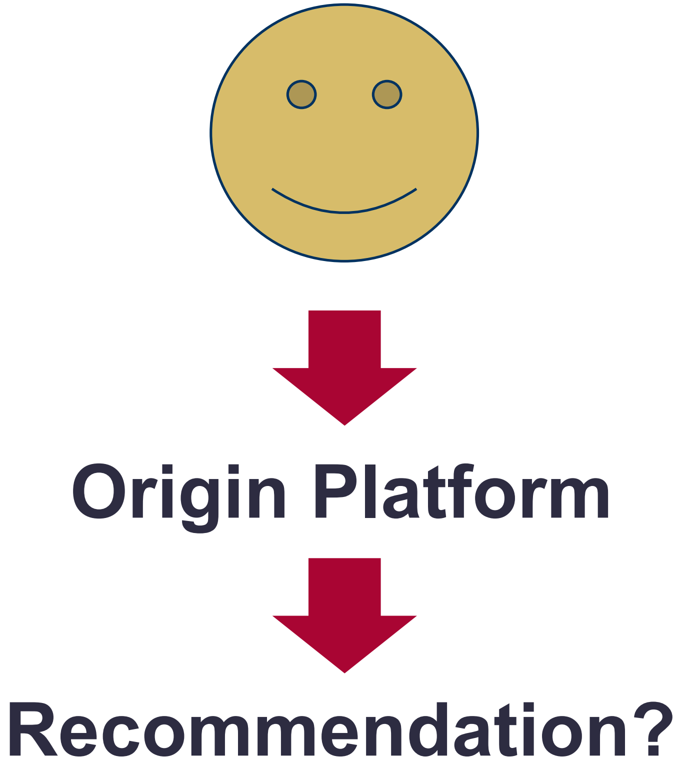
User 1: Game 1 > 2 > 5 > 3 > …
User 3: Game 5 > 7 > 2 > 3 > …

The Problem: Researcher’s Perspective

The goal: For any player, understand his/her preferences over all possible games

Key challenges:

- Providing EA *useful outputs*
- The *cold start problem*

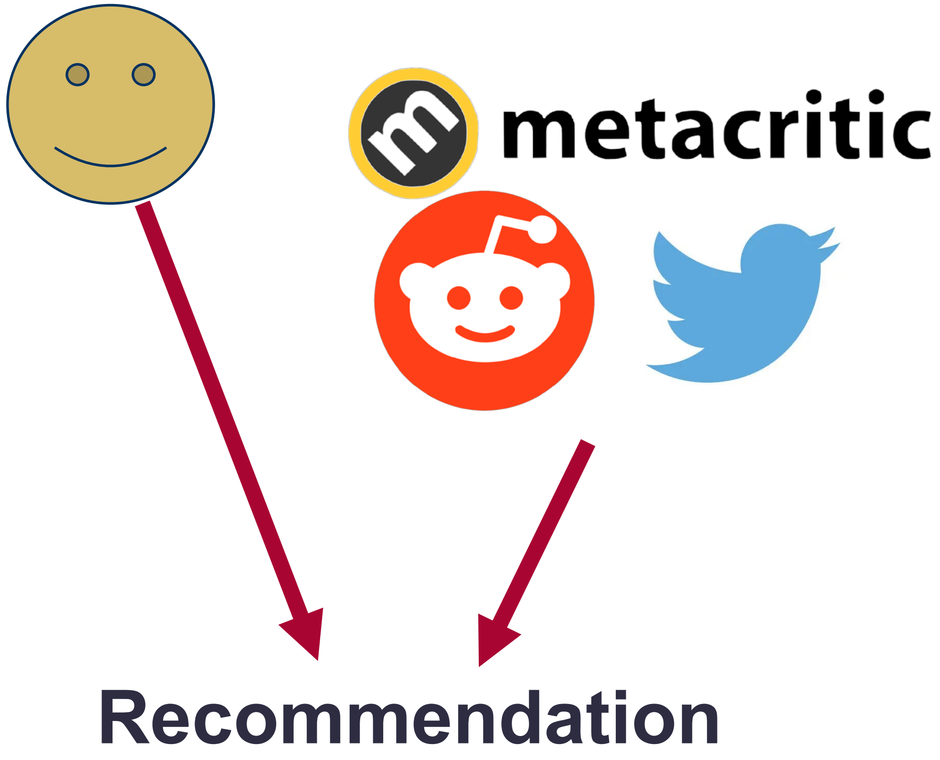


The Problem: Researcher’s Perspective

The goal: For any player, understand his/her preferences over all possible games

Key challenges:

- Providing EA *useful outputs*
- The *cold start problem*
- Learning from *diverse data sources*



The Problem: Researcher’s Perspective

The goal: For any player, understand his/her preferences over all possible games

Key challenges:

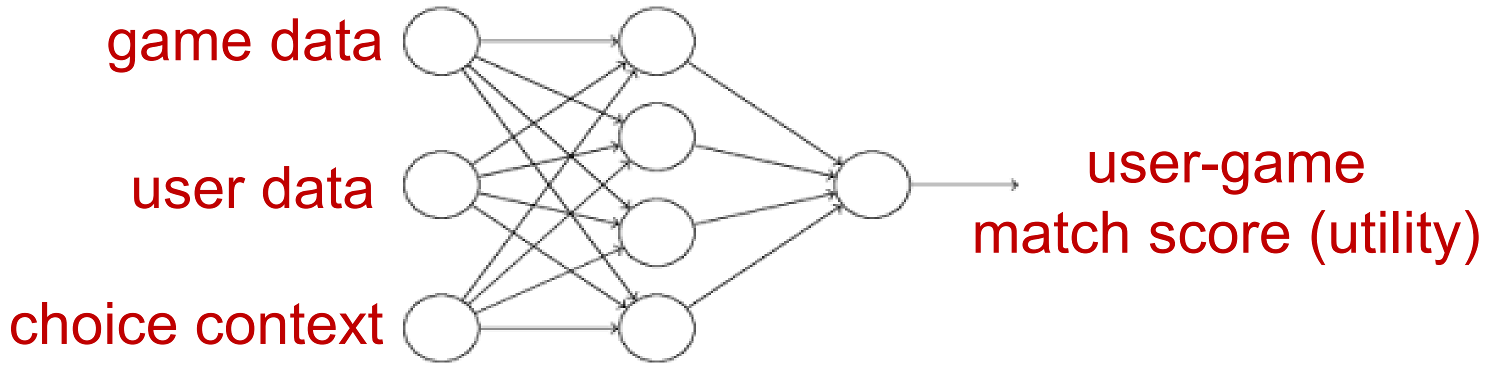
- Providing EA *useful outputs*
- The *cold start problem*
- Learning from *diverse data sources*
- Recommendation *variety*

Our Approach: User-Game Match Scores

user plays game X at time t → **user chooses game X over any other game at time t** → **user's utility from game X is higher than utility from any other game at time t**

How do we estimate user U's utility for game X at time t?

We fuse *game data*, *user data*, and *contextual data* via neural networks:



Details

- **What data goes into this network?**
  - **User history** – users are represented as the average of their past 20 games
  - **Recency** and **Frequency** of game play
  - **Metacritic data** – numeric ratings, temporal information, *and the text as word embeddings*
- **How do we learn the parameters of the network?**
  Minimize the *Bayesian personalized ranking (BPR)* loss:

Probability of choosing the wrong game

$$\text{Loss} = E\left[\frac{\exp(N)}{\exp(N) + \exp(S)}\right] + \text{Regularization}(\alpha)$$

where S is the user's score for the game the user played, and N is the score for a randomly selected other game.

Key Insights

Evaluating success: **mean reciprocal rank**

- Model predicts game is rank 1 → user plays that game → MRR = 1/1 = 1
- Model predicts game is rank 2 → user plays that game → MRR = 1/2 = 0.5

Simple Embedding-Only Baselines

+

Avg. Game History Embedding

+

Explicit Recency-Frequency Metrics

+ Metacritic

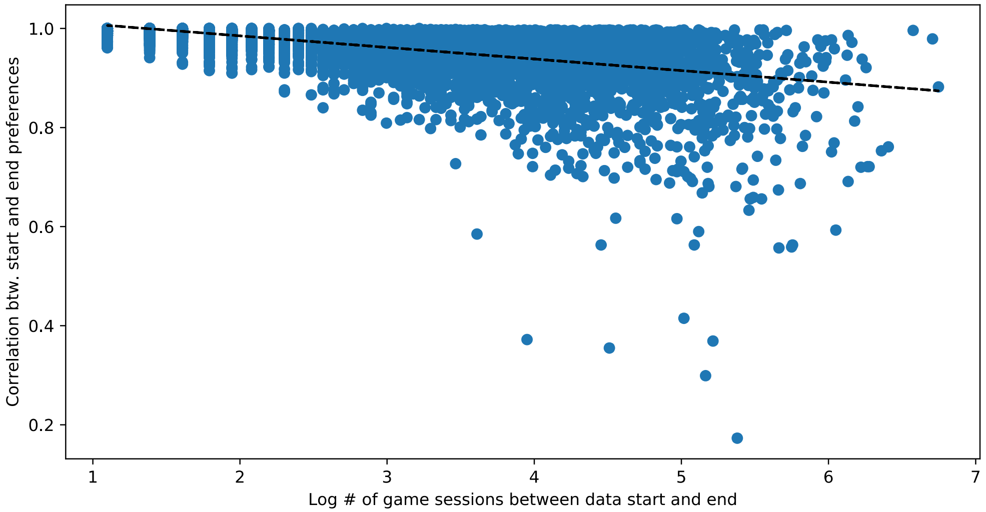
	MRR Test	New Users	New Games
Random	0.05	0.05	0.05
Model1	0.34	0.36	0.02
Model2	0.30	0.20	0.02
Model3	0.60	0.64	0.04
Model3a	0.59	0.63	0.01
Model3b	0.33	0.35	0.02
Model3c	0.49	0.27	0.02
Model3 stochastic	0.50	0.53	0.02
Model4	0.72	0.75	0.47
Model4a	0.72	0.75	0.49
*Model5	0.71	0.74	0.45
*Model6	0.71	0.74	0.47

What does this mean?

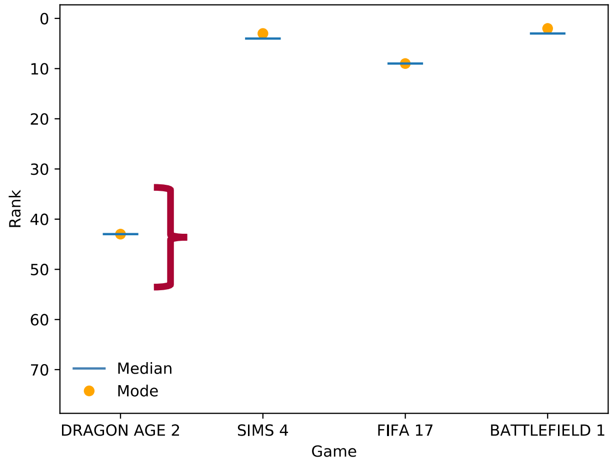
- Past behavior data seems to be the most informative type of data.
- Adding explicit game-specific recency-frequency really helps
- Context covariates help for handling *cold start problems*

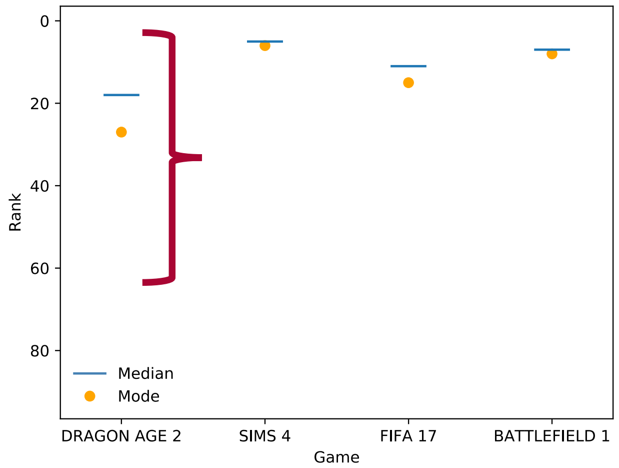
Variance and Variety of Recommendations

People's preferences do change over time (i.e. dynamics matter)



Adding stochasticity *may* lead to higher variety of recommendations (Quasi-Bayesian approach to embeddings)





Academic Contributions

- Paper based on this data and findings is a work-in-progress

- Main contribution: dynamic, adaptive, contextual recommendations
  - Leveraging deep learning + BPR for preference modeling
  - Data fusion: mixing wide variety of data sources for single, coherent ranking
  - Stochasticity for explore + exploit learning