OPFY on CSR

Andrew Stanton
Mohit Nayyar

Thanks to:
Diane Hu
Aakash Sabarwal
Devin Guillory

#csr

Agenda

1. Old OPFY

2. New OPFY

3. Wins on new OPFY

4. CSR in OPFY context

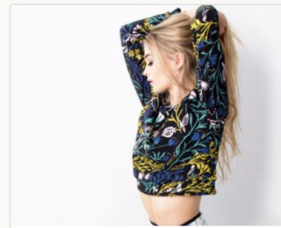5. OPFY CSR connections

6. CSR Wins

Old OPFY

Our picks for Lisa

Removable Wallpaper // Avoc...	Golden Cactus Big Tee, overs...	Indigo Blue Bandana - hand d...	Cropped Sweatshirt - Hand P...	Cactus Bralette Underwear S...	Evil Eye Earrings, Evil Eye, Ro...
KateZarembaCompa... $36.00	SimkaSol $72.00	APPRVL $25.00	thiefandbandit $110.00	thiefandbandit $82.00	RabbitsFantasyWorld $16.27

1. On homepage for signed-in users
2. Arizona dataset:
   a. Key: user id
   b. Value: listing ids reverse sorted by score on trained model
   c. Model: logistic regression trained on clicks using vowel wabbit library

New OPFY

- New Arizona dataset

  - Key: user id

  - Value: Set(listing_id1 -> features of listing_id1, …)

- When user visits etsy homepage:

  - Score each listing

  - Reverse sort by rank and send back to etsyweb

Wins on new OPFY

- Older OPFY would calculate recommendations for all

  our users

- Newer OPFY only for users that come to the website

- Saves time and resources on hadoop

- In real-time, bro

CSR in OPFY Context
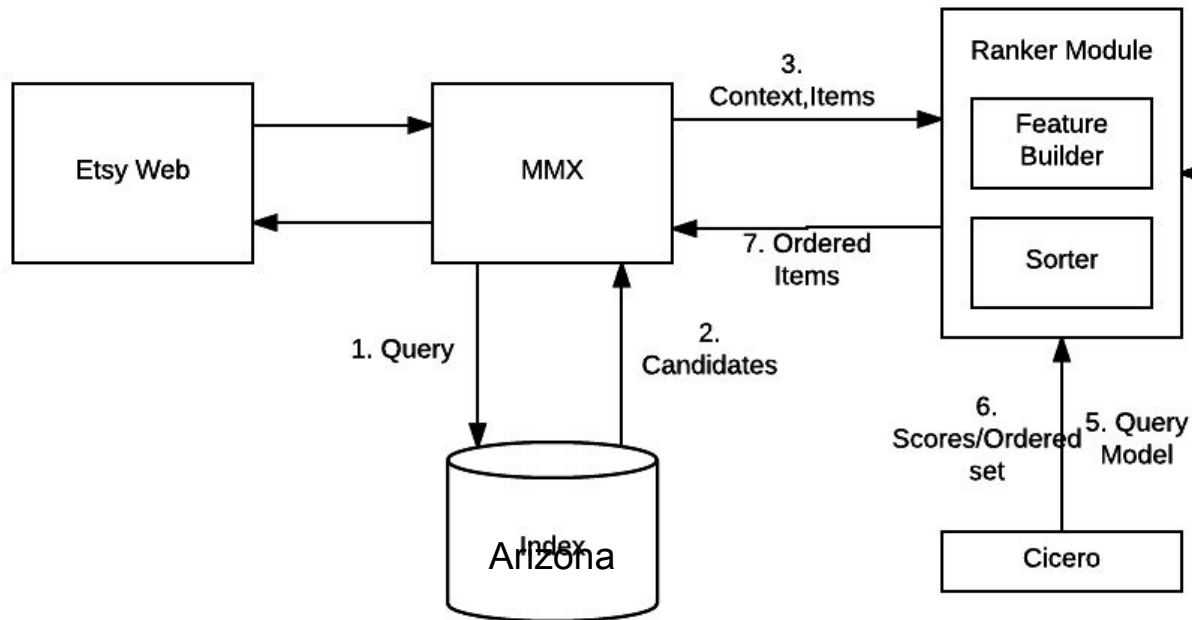
This part not used by OPFY

- Feature builder not used by OPFY
- Reranker Module currently part of MMX

CSR in OPFY Context (contd.)

- OPFY training happens on BigData/Computron nightly in oozie workflow

- The trained modules are deployed to Cicero as part of the workflow

CSR Wins

- Can use real-time context (Yuuuge!)

- Recommendations have a 2-day lag, we can make

  them relevant to the current visit, for instance.