

Rank Model Feature

Monday, April 1, 2024

11:39 AM

Model Features

- ① user profile
- ② item profile
- ③ user history feature
- ④ item history feature
- ⑤ scenario feature

User Profile

- ① user ID (embedding as 32 or 64 vector)
- ② user person info: gender, age group.
- ③ account info: new, old, active, inactive
- ④ interested category: sport? movie?

Item Profile:

- ① item ID (embedding)
- ② item "age"
- ③ item published location (geo hash, city)
- ④ item title, key word, description
- ⑤ item property: length, quality, "elegant"

Rank Model Feature

Monday, April 1, 2024

11:39 AM

User History Info:

(30, 7, 3, 1)

- ① user exposure, click, like... in last N days
- ② user history data group by item category

Item History Info:

- ① item being clicked, liked, ... in last N days
- ② item history info grouped by user age, gender, etc.
- ③ item's author info:
(a) # of published item (b) # of follower (c) # of click...

Scenario Feature:

- ① user location
- ② current time being
- ③ is holiday or any special time
- ④ different login device

Rank Model Feature

Monday, April 1, 2024

11:39 AM

Feature Engineering Process:

Discrete Feature \rightarrow embedding

e.g. user ID, item ID, item category, location

Continuous Feature \rightarrow group by bin

e.g. age group, video length, item description length

Continuous Feature \rightarrow other transform

e.g. # of clicks $\rightarrow \log(1+x)$

of like \rightarrow percent of like \rightarrow smooth

Feature Coverage:

- ① Users choose NOT to provide some info: age, gender...
- ② if feature is missing, experiment with default value
- ③ ideally, make feature cover 100% samples.

Rank Model Feature

Monday, April 1, 2024

11:39 AM

Online Service :

