Recommender Systems

Machine Learning for the Web

a talk for Python Sevilla

Seville, February 1, 2017

$> whoami

José M. Camacho

ML developer at geographica

jcamacho@geographica.gs

@josemazo

josemazo

MACHINE LEARNING

CLASSIFICATION

REGRESSION

CLUSTERING

RECOMMENDATION

GENERATION

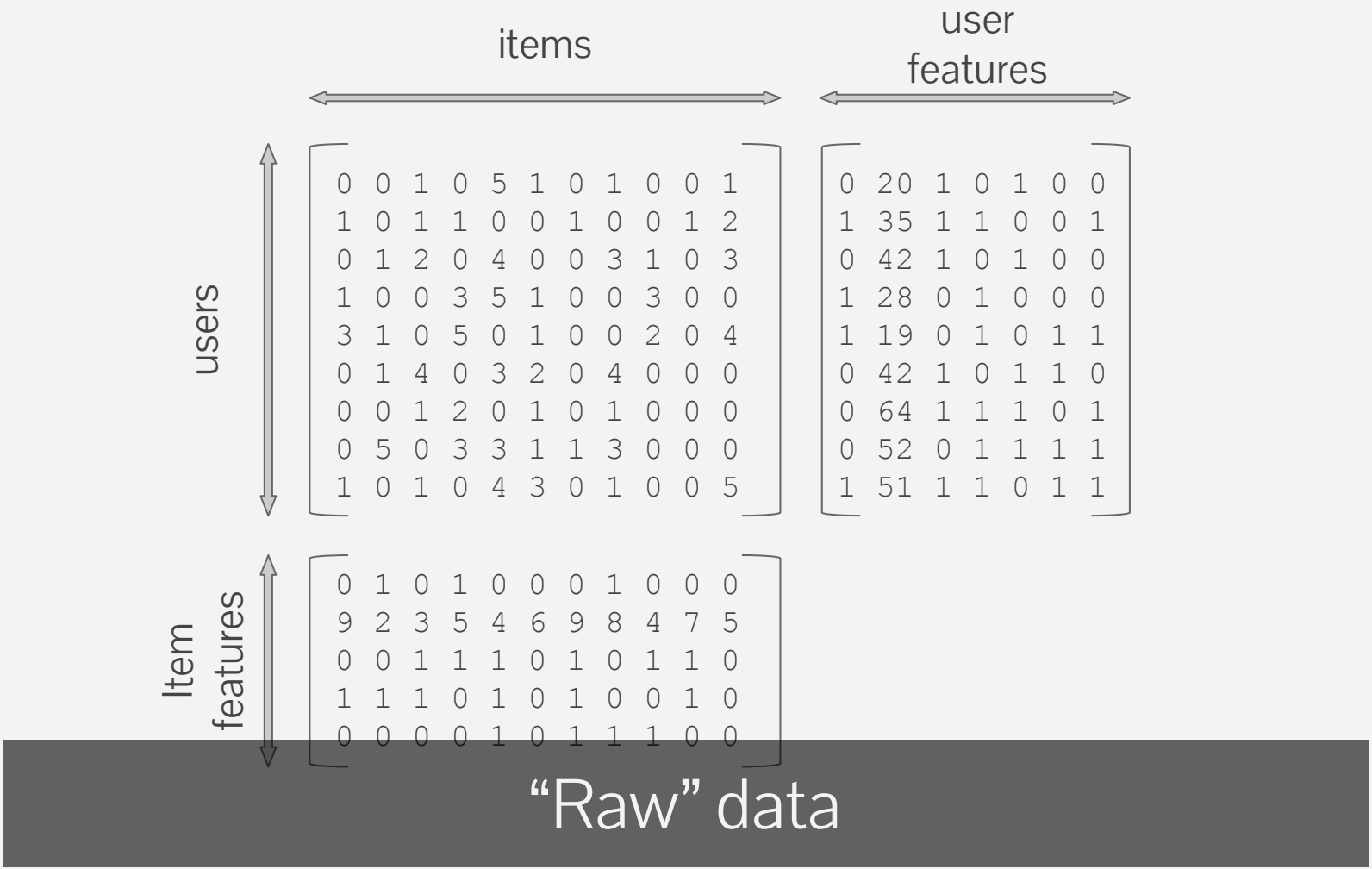

What do you recommend me?

COLLABORATIVE FILTERING

CONTENT-BASED FILTERING

HYBRID SYSTEMS

FAKE HYBRID SYSTEMS


items
user features
users
0 0 1 0 5 1 0 1 0 0 1
1 0 1 1 0 0 1 0 0 1 2
0 1 2 0 4 0 0 3 1 0 3
1 0 0 3 5 1 0 0 3 0 0
3 1 0 5 0 1 0 0 2 0 4
0 1 4 0 3 2 0 4 0 0 0
0 0 1 2 0 1 0 1 0 0 0
0 5 0 3 3 1 1 3 0 0 0
1 0 1 0 4 3 0 1 0 0 5
0 20 1 0 1 0 0
1 35 1 1 0 0 1
0 42 1 0 1 0 0
1 28 0 1 0 0 0
1 19 0 1 0 1 1
0 42 1 0 1 1 0
0 64 1 1 1 0 1
0 52 0 1 1 1 1
1 51 1 1 0 1 1
Item features
0 1 0 1 0 0 0 1 0 0 0
9 2 3 5 4 6 9 8 4 7 5
0 0 1 1 1 0 1 0 1 1 0
1 1 1 0 1 0 1 0 0 1 0
0 0 0 0 1 0 1 1 1 0 0
“Raw” data

$$\hat{r}_{ui} = f(q_u \cdot p_i + b_u + b_i)$$

$$q_u = \sum_{j \in f_u} e_j^U \qquad p_i = \sum_{j \in f_i} e_j^I \qquad b_u = \sum_{j \in f_u} b_j^U \qquad b_i = \sum_{j \in f_i} b_j^I$$

Latent factor models

Maciej Kula (Lyst)'s LightFM

Python Sevilla repository

QUESTIONS

PLEASE!

A MILLION

THANKS

FOR BEING

HERE TODAY