Machine Learning at ColorTV

Marcin Ciura

2016

Two recommendation systems

- Ad serving (collaborative filtering)
- Content recommendation (content-based filtering)

Collaborative filtering (1)

	Nowak	Kowalska	Wiśniewski
50 twarzy Greya	1		
Jadłonomia			1
Moje wypieki		1	1

Collaborative filtering (2)

$$\begin{array}{l} \\ \text{50 twarzy Greya} \\ \text{Jadłonomia} \\ \text{Moje wypieki} \end{array} \begin{array}{c} \begin{array}{cc} \mathsf{F}_1 & \mathsf{F}_2 \end{array} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \end{array} \times \begin{array}{c} \\ \mathsf{F}_1 \\ \mathsf{F}_2 \end{array} \begin{array}{c} \begin{array}{ccc} \text{Nowak} & \text{Kowalska} & \text{Wi sniewski} \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \end{array} = \begin{array}{l} \\ \text{50 twarzy Greya} \\ \text{Jadłonomia} \\ \text{Moje wypieki} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Nowak} & \text{Kowalska} & \text{Wi sniewski} \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \end{array}$$

Collaborative filtering (3)

Fill zero elements of a given matrix M, factorizing it as a product of matrices P and Q^T of a given rank r, minimizing the approximation error $\|M - PQ^T\|$ calculated for non-zero elements of matrix M.

Collaborative filtering (4)

Practical considerations:

- Factorize offline with Apache Spark
- Obey ad and campaign constraints
- Demote already seen ads