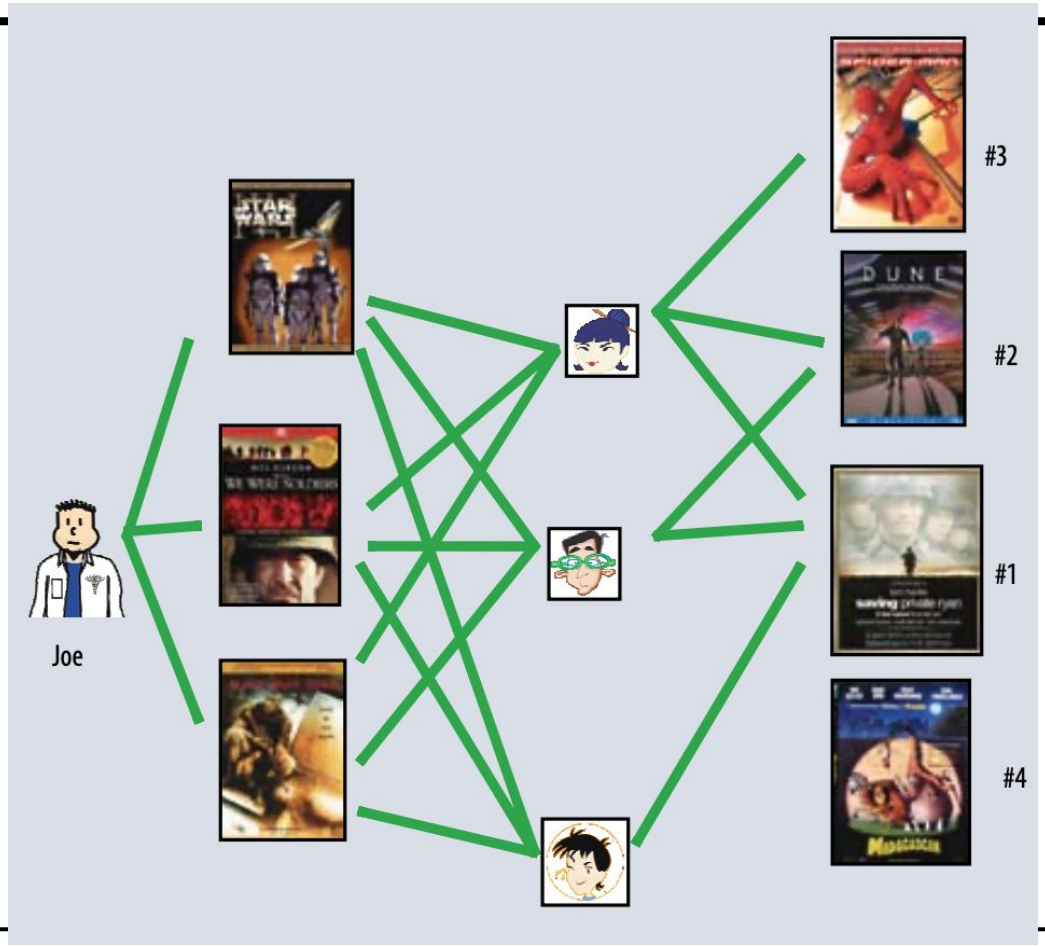
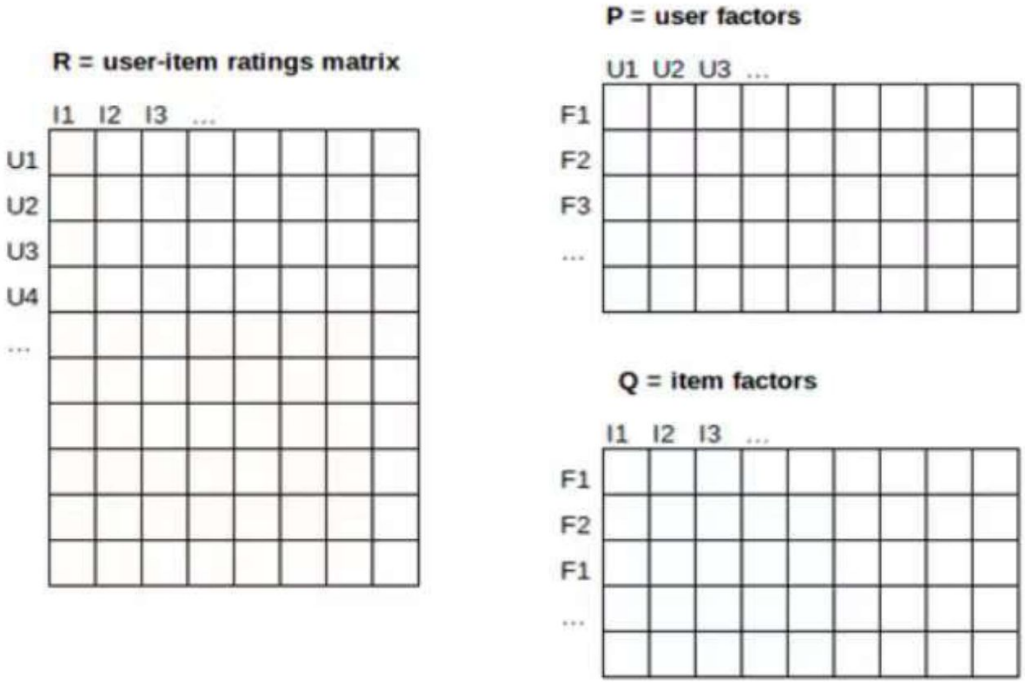
Recommender Systems

Group 01:
Team Members:
1. Chen, Shenghan sc4615@columbia.edu
2. Chen, Shuyi sc4489@columbia.edu
3. Guo, Junyan jg4184@columbia.edu
4. Heagy, David dh2868@columbia.edu
5. Xie, Wenjie wx2223@columbia.edu

The designated presenter is Chen, Shuyi


Joe
#3
#2
#1
#4

Matrix factorization



Assigned pairings

-factorization algorithm: Stochastic Gradient Descent(A1)

- regularization: Penalty of Magnitudes (R1)+ Bias and Interecepts(R2) vs Penalty of Magnitudes + Temporal Dynamics(R3)

- Postpocessing:KNN(P3)

Dataset

610 users

9724 ratings

Models

Our Default Model: Stochastic Gradient Descent

$$min_{q*p*} \sum_{(u,i) \in K} (r_{ui} - q_i^T p_u)^2 + \lambda(||q_i||^2 + ||p_u||^2)$$

Models with Regularization

Predicted ratings for adding bias

$$\hat{r}_{ui} = \mu + b_i + b_u + q_i^T p_u$$

Predicted ratings for adding Temporal Dynamics

$$\hat{r}_{ui}(t) = q_i^T p_u + \mu + b_i + b_u + \alpha_u dev_u(t)$$

Notation	Definition
p_u	Vector associated with user u
q_i	Vector associate with iterm i
b_i	Observed bias of item i
b_u	Observed bias of user u
u	Global average rating
r_i	Observed rating
$\hat{r}_x$	Predicted rating
f	Dimension of factor
λ	Penalty parameter

Postprocessing SVD with KNN

Linear Regression

Linear Regression Function

$$\widehat{r_{ui}} = \beta_0 + \beta_1 * p_u q^t{}_i + \beta_2 * b_u + \beta_3 * b_i + \beta_4 * knn$$

$$\widehat{r_{ui}} = \beta_0 + \beta_1 * p_u q^t{}_i + \beta_2 * b_u + \beta_3 * b_i + \beta_4 * a_u dev_u(t) + \beta_5 * knn$$

	A1+R1+R3+P2	A1+R1+R2+P2
bi	0.88	0.89
bu	0.83	0.81
pq	1.48	1.14
dev	0.90	
knn	-0.01	0.01

Result Summary

RMSE	Train Data	Test Data
A1+P2	1.01	1.16
A1+R1R2+P2	1.20	1.20
A1+R1R3+P2	1.28	1.32
Linear Regression: A1+R1R2+P2	0.85	0.74
Linear Regression: A1+R1R3+P2	0.72	0.93