

Course Update *Out*

Current Plan (updated) *Due*

- HW 8 (online)
- Mini-project proposal
- HW 9 (online)
- HW 10 (written/prog)
- Midterm 2
- Mini-project

Sun	Mon	Tue	Wed	Thu	Fri	Sat
2	3	4	5 HW8 Proj	6	7	8
9 HW8	10 HW9 HW10	11	12	13	14	15
16	17 HW9	18	19 Prop	20	21	22 HW10
23	24	25	26 MT2	27	28	29
30	1	2	3	4	5 Proj	6

# Plan

## Last time

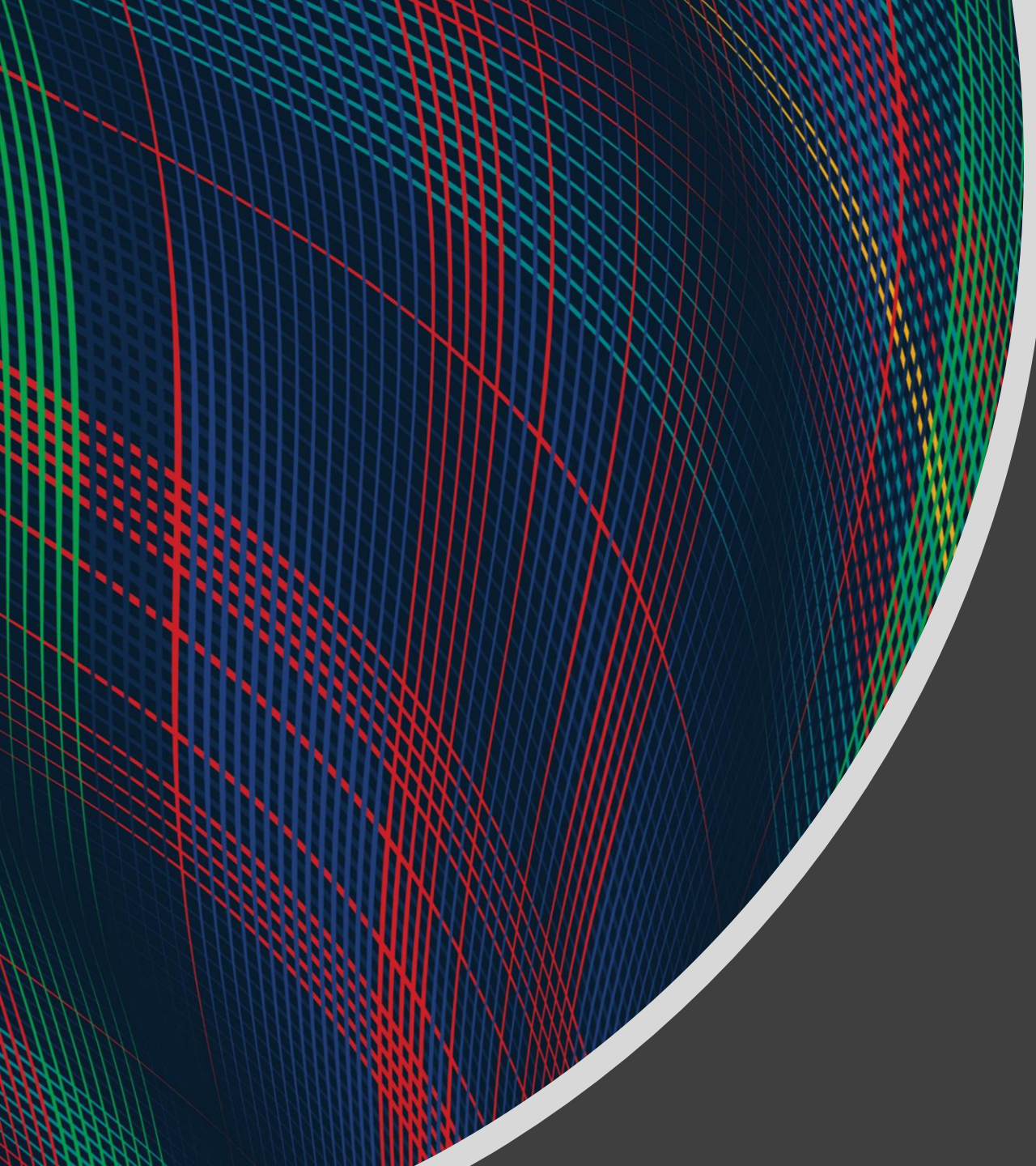
- Unsupervised Learning: Dimensionality Reduction

## Today

- Recommender Systems
- Unsupervised Learning: Clustering
  - K-means

## Next time

- Unsupervised Learning: Clustering
  - Gaussian mixture models and expectation maximization



10-315

Introduction to ML

Recommender Systems

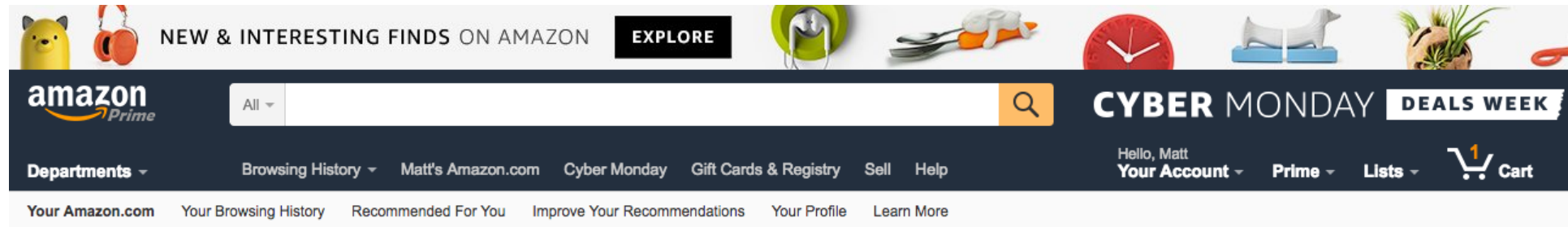
Instructor: Pat Virtue

# Recommender Systems

## A Common Challenge:

- Assume you're a company selling items of some sort: movies, songs, products, etc.
- Company collects millions of ratings from users of their items
- To maximize profit / user happiness, you want to recommend items that users are likely to want

# Recommender Systems



NEW & INTERESTING FINDS ON AMAZON EXPLORE

amazon Prime

All

CYBER MONDAY DEALS WEEK

Departments

Browsing History

Matt's Amazon.com

Cyber Monday

Gift Cards & Registry

Sell

Help

Hello, Matt

Your Account

Prime

Lists

Cart

Your Amazon.com

Your Browsing History

Recommended For You

Improve Your Recommendations

Your Profile

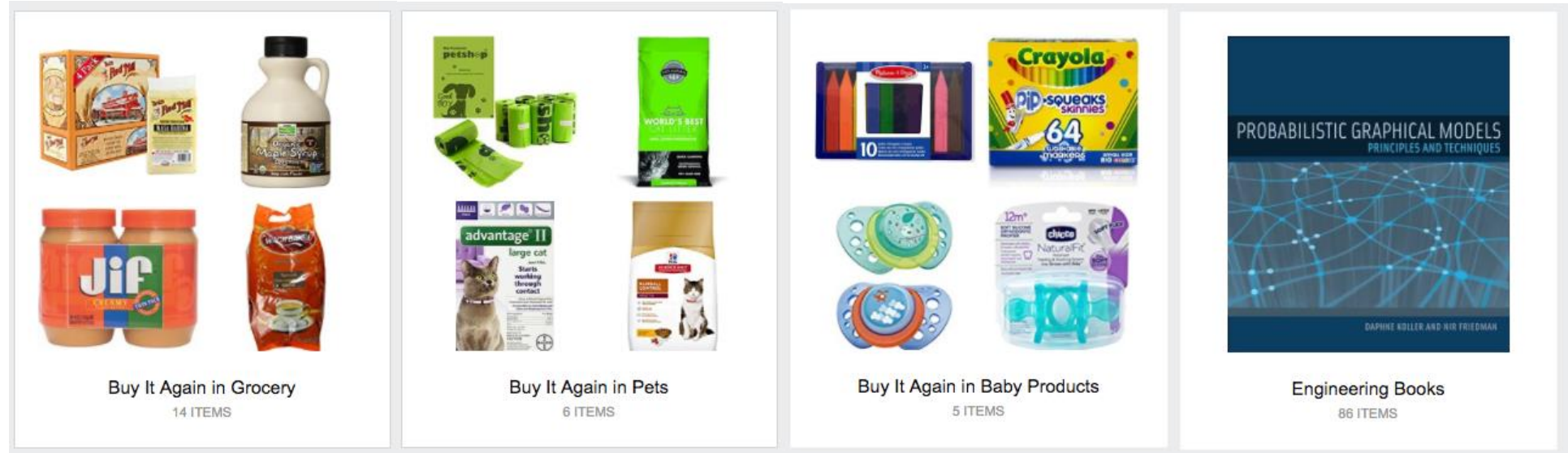
Learn More



You could be seeing useful stuff here!  
Sign in to get your order status, balances and rewards.

Sign In

## Recommended for you, Matt



Buy It Again in Grocery  
14 ITEMS

Buy It Again in Pets  
6 ITEMS

Buy It Again in Baby Products  
5 ITEMS

Engineering Books  
86 ITEMS



# Recommender Systems

**NETFLIX**

## Netflix Prize

**COMPLETED**

Home Rules Leaderboard Update

### Congratulations!

The Netflix Prize sought to substantially improve the accuracy of predictions about how much someone is going to enjoy a movie based on their movie preferences.

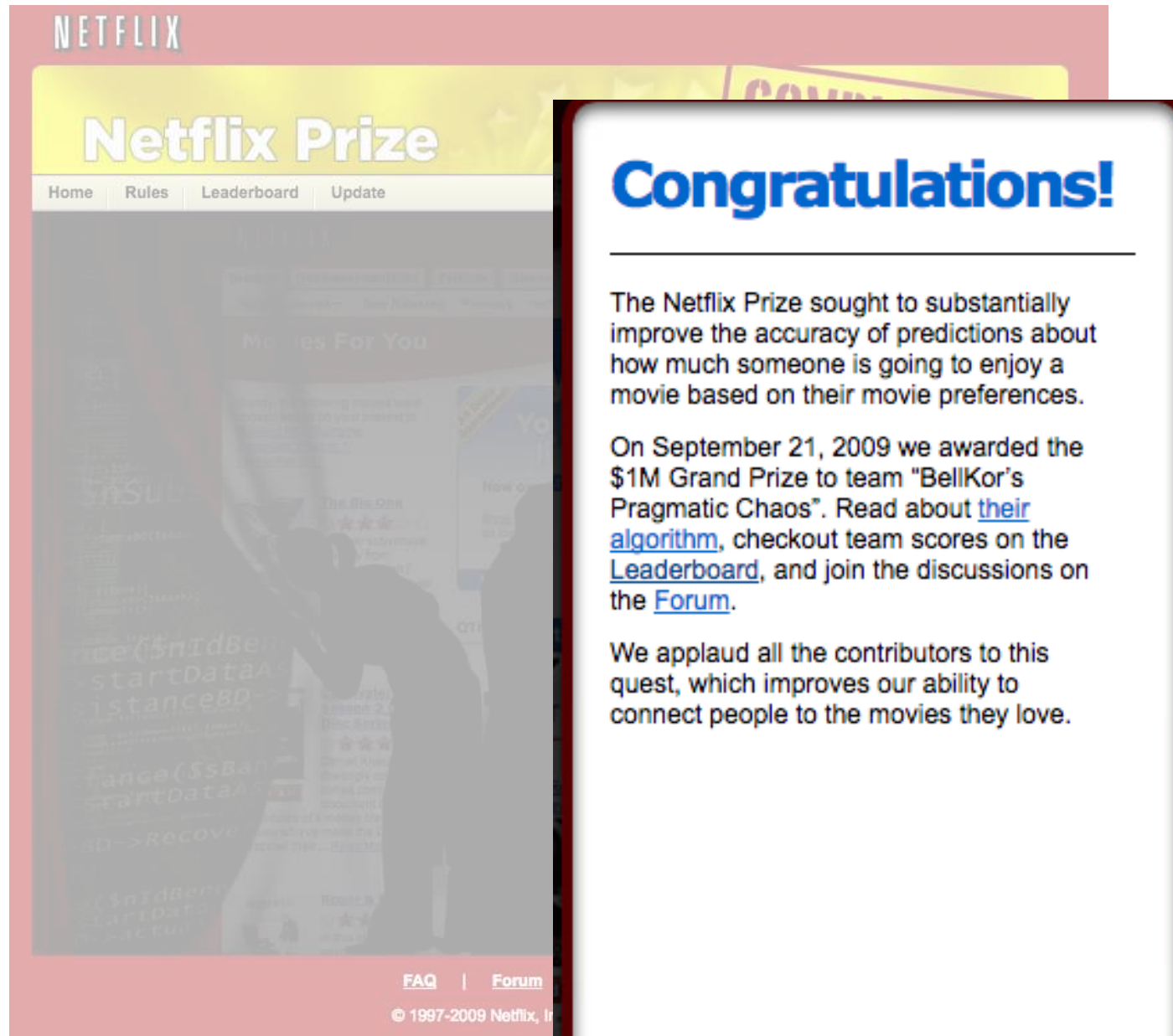
On September 21, 2009 we awarded the \$1M Grand Prize to team "BellKor's Pragmatic Chaos". Read about [their algorithm](#), checkout team scores on the [Leaderboard](#), and join the discussions on the [Forum](#).

We applaud all the contributors to this quest, which improves our ability to connect people to the movies they love.

FAQ | Forum | Netflix Home

© 1997-2009 Netflix, Inc. All rights reserved.

# Recommender Systems



The image shows a screenshot of the Netflix Prize website. The top navigation bar includes 'Home', 'Rules', 'Leaderboard', and 'Update'. The main content area is titled 'Movies For You' and features a grid of movie recommendations. A large, semi-transparent text box is overlaid on the right side of the page, containing the following text:

## Congratulations!

The Netflix Prize sought to substantially improve the accuracy of predictions about how much someone is going to enjoy a movie based on their movie preferences.

On September 21, 2009 we awarded the \$1M Grand Prize to team "BellKor's Pragmatic Chaos". Read about [their algorithm](#), checkout team scores on the [Leaderboard](#), and join the discussions on the [Forum](#).

We applaud all the contributors to this quest, which improves our ability to connect people to the movies they love.

FAQ | Forum

© 1997-2009 Netflix, Inc.

# AI System Design: Movie Recommendation

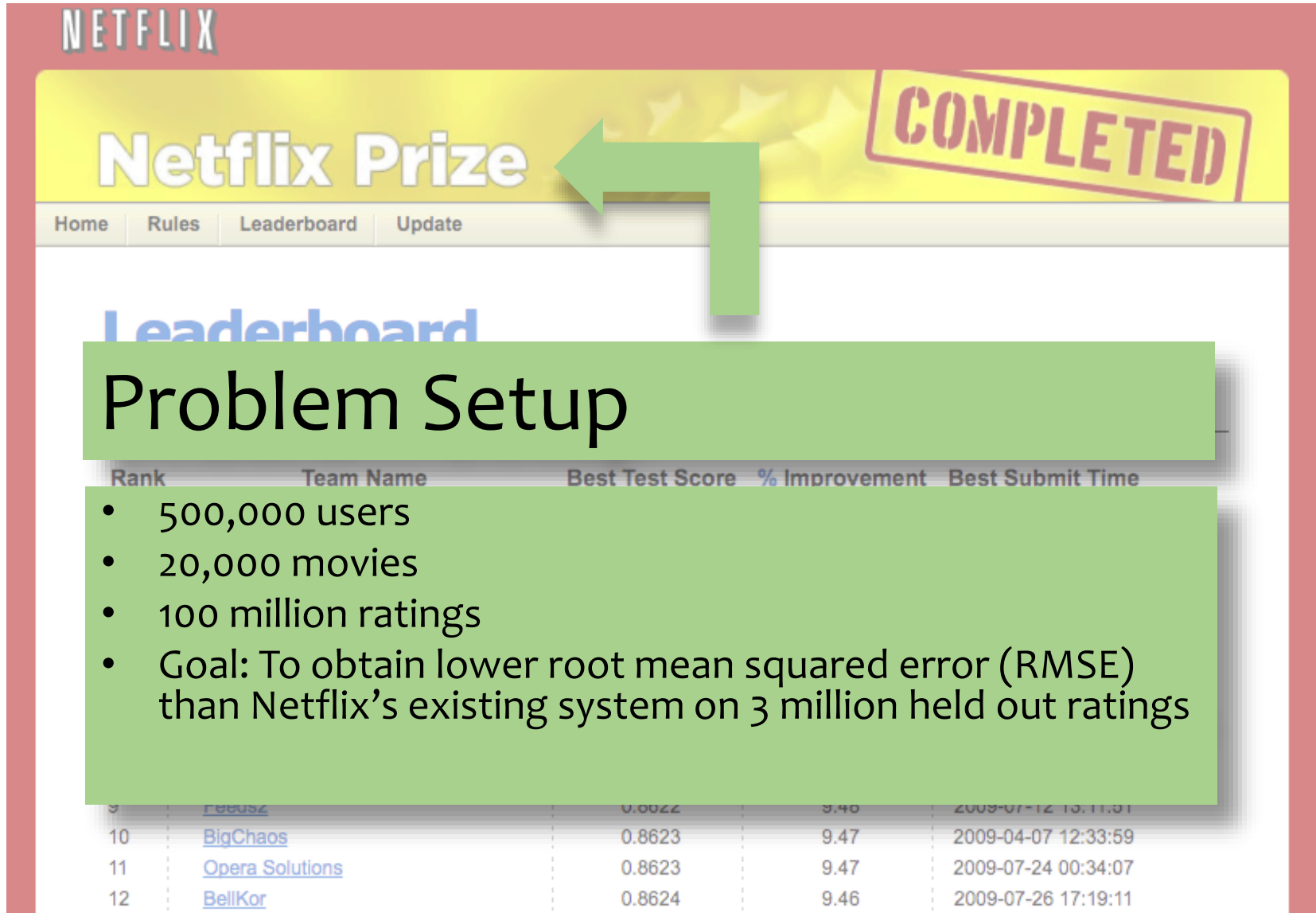
Task

Experience

Performance measure



# Recommender Systems



NETFLIX

## Netflix Prize

Home Rules Leaderboard Update

### Leaderboard

**Problem Setup**

- 500,000 users
- 20,000 movies
- 100 million ratings
- Goal: To obtain lower root mean squared error (RMSE) than Netflix's existing system on 3 million held out ratings

Rank	Team Name	Best Test Score	% Improvement	Best Submit Time
9	<a href="#">Feus2</a>	0.8622	9.48	2009-07-12 15:11:01
10	<a href="#">BigChaos</a>	0.8623	9.47	2009-04-07 12:33:59
11	<a href="#">Opera Solutions</a>	0.8623	9.47	2009-07-24 00:34:07
12	<a href="#">BellKor</a>	0.8624	9.46	2009-07-26 17:19:11

# ML System Design: Movie Recommendation

## Model

# ML System Design: Movie Recommendation

## Model

# Recommender Systems

## Setup:

- Items:  
movies, songs, products, etc.  
(often many thousands)
- Users:  
watchers, listeners, purchasers, etc.  
(often many millions)
- Feedback:  
5-star ratings, not-clicking 'next', purchases,  
etc.

## Challenge:

- Users only rate a small number of items  
(the user/item rating data is sparse)

	Doctor Strange	Star Trek: Beyond	Zootopia
Alita	1		5
BB-8	3	4	
C-3PO	3	5	2

# Different Approaches

## Item-based (*Content filtering*)

- Features about each item
- Given an item, other “close” items have similar values
- e.g. Pandora.com, music genome project



# Different Approaches

## Item-based (*Content filtering*)

- Features about each item
- Given an item, other “close” items have similar values
- e.g. Pandora.com, music genome project

## User-based

- Features about each user
- Given a user, other “close” users have similar preferences
- *Market segmentation*

## Learning user-item relationship

- Can be done without features on either user or item
- Collaborative filtering techniques

# Collaborative Filtering

# Collaborative Filtering

## Everyday Examples of Collaborative Filtering...

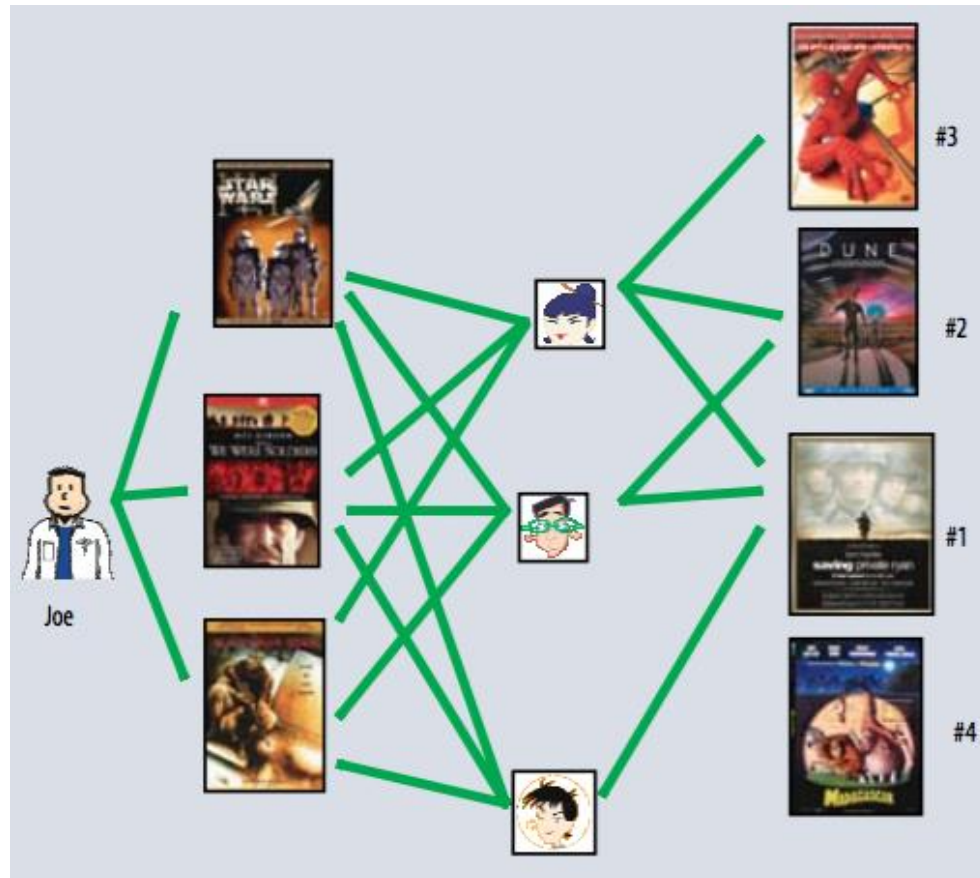
- Bestseller lists
- Top 40 music lists
- The “recent returns” shelf at the library
- Unmarked but well-used paths thru the woods
- The printer room at work
- “Read any good books lately?”
- ...

## Common insight: personal tastes are correlated

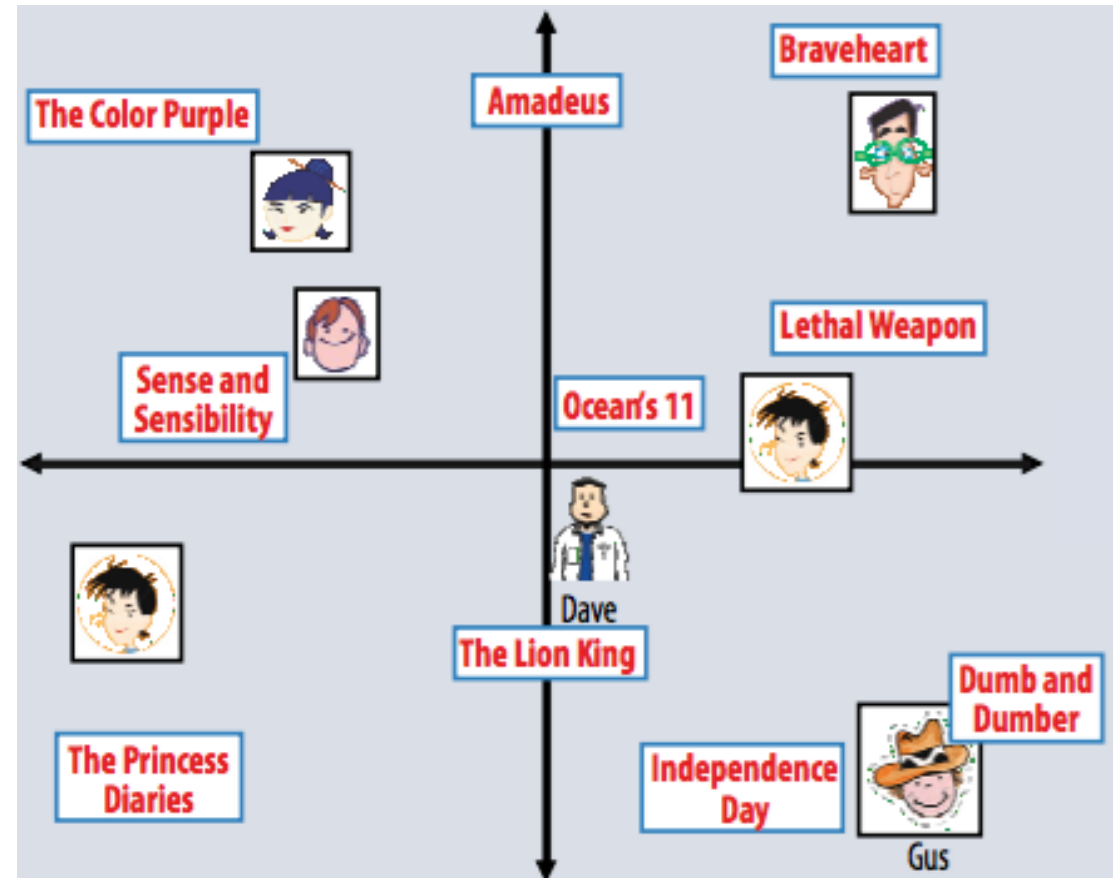
- If Alita and BB-8 both like X and Alita likes Y then BB-8 is more likely to like Y
- especially (perhaps) if BB-8 knows Alita

# Two Types of Collaborative Filtering

## 1. Neighborhood Methods

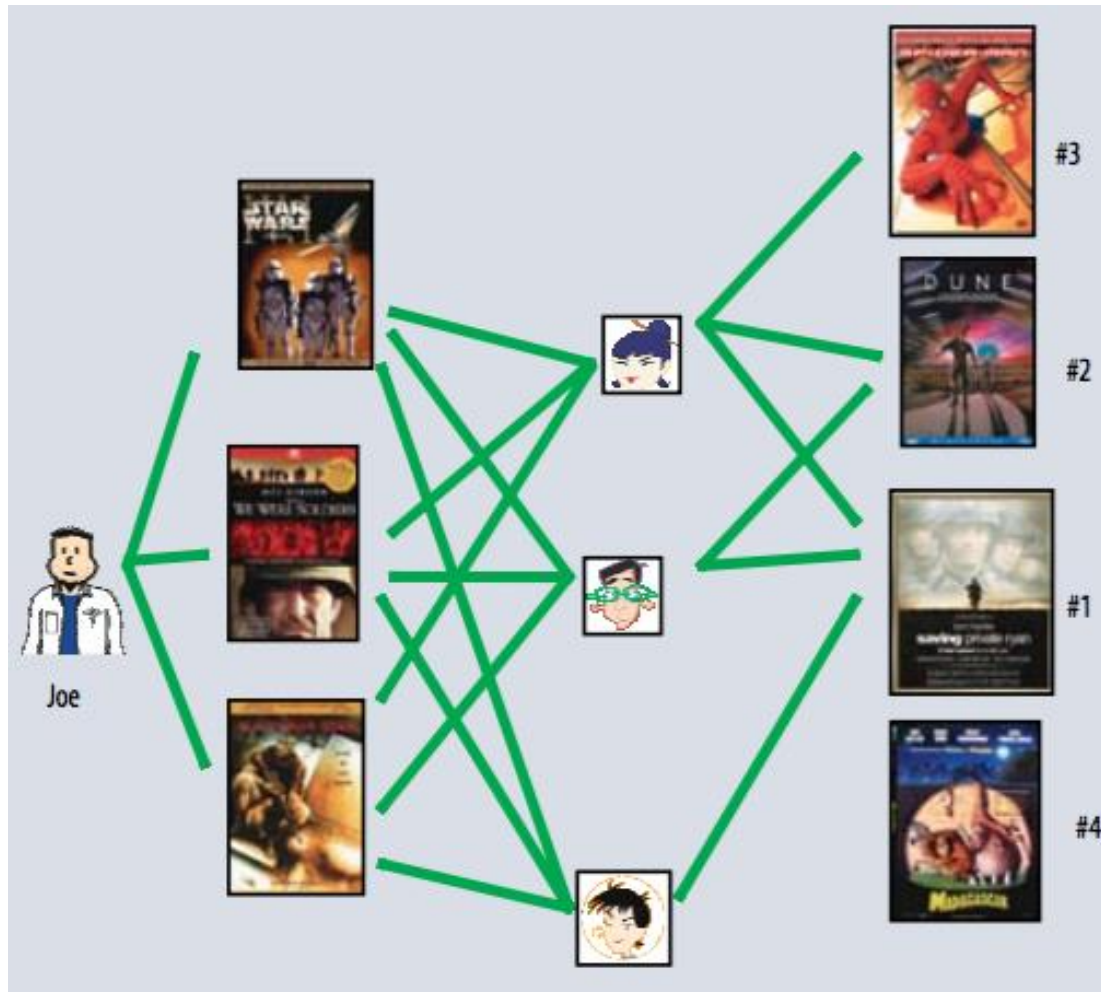


## 2. Latent Factor Methods



# Two Types of Collaborative Filtering

## 1. Neighborhood Methods



In the figure, assume that a green line indicates the movie was **watched**

### Algorithm:

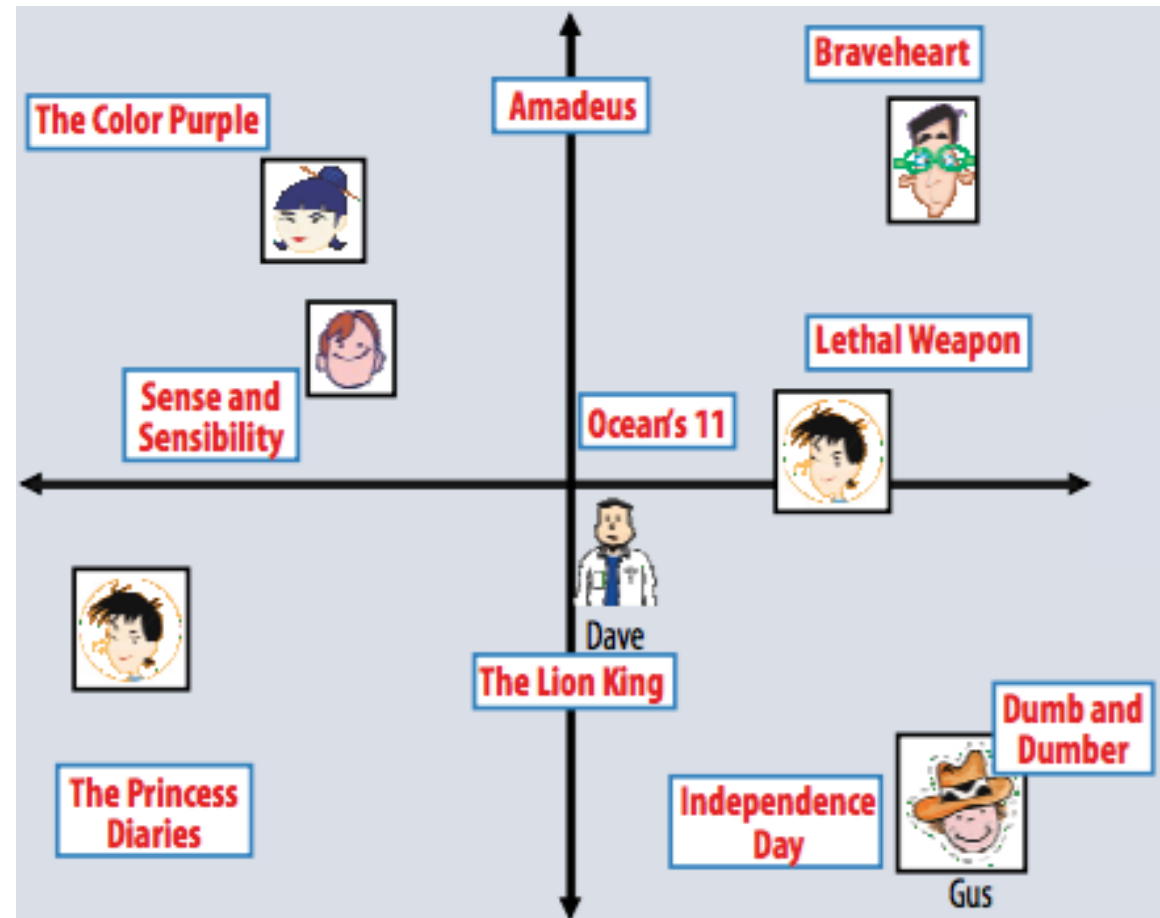
1. **Find neighbors** based on similarity of movie preferences
2. **Recommend** movies that those neighbors watched



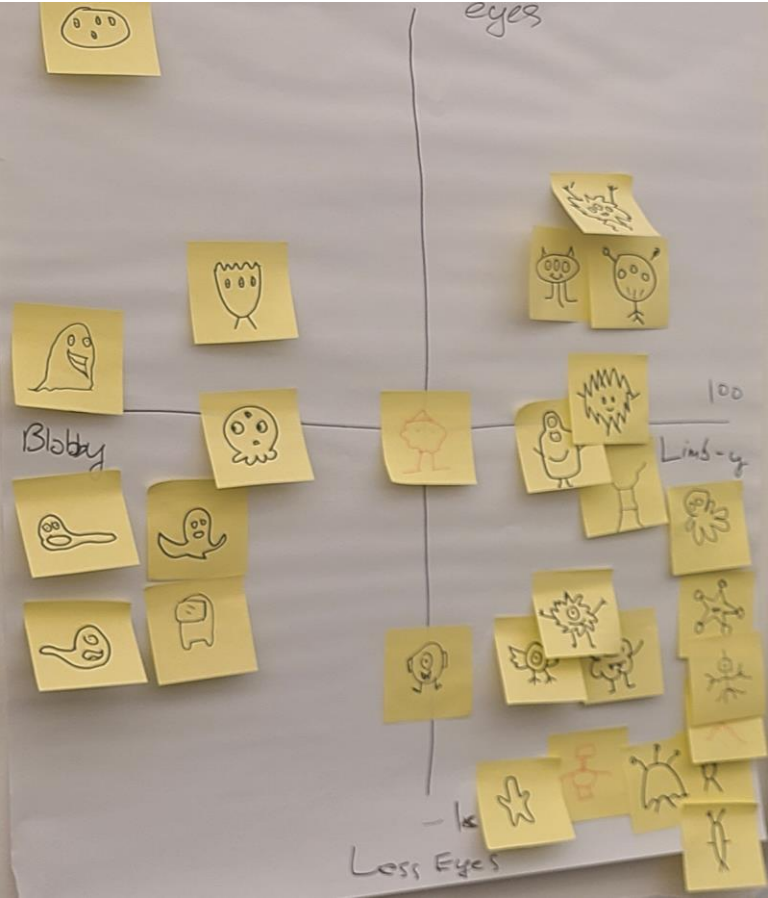
# Two Types of Collaborative Filtering

## 2. Latent Factor Methods

- Assume that both movies and users can be mapped to the same **feature space**
- **Recommend** a movie based on its **proximity** to the user in the feature (latent) space
- **Example algorithm:** Matrix Factorization



# Background: Learn a Feature Space



# Background: Learn a Feature Space

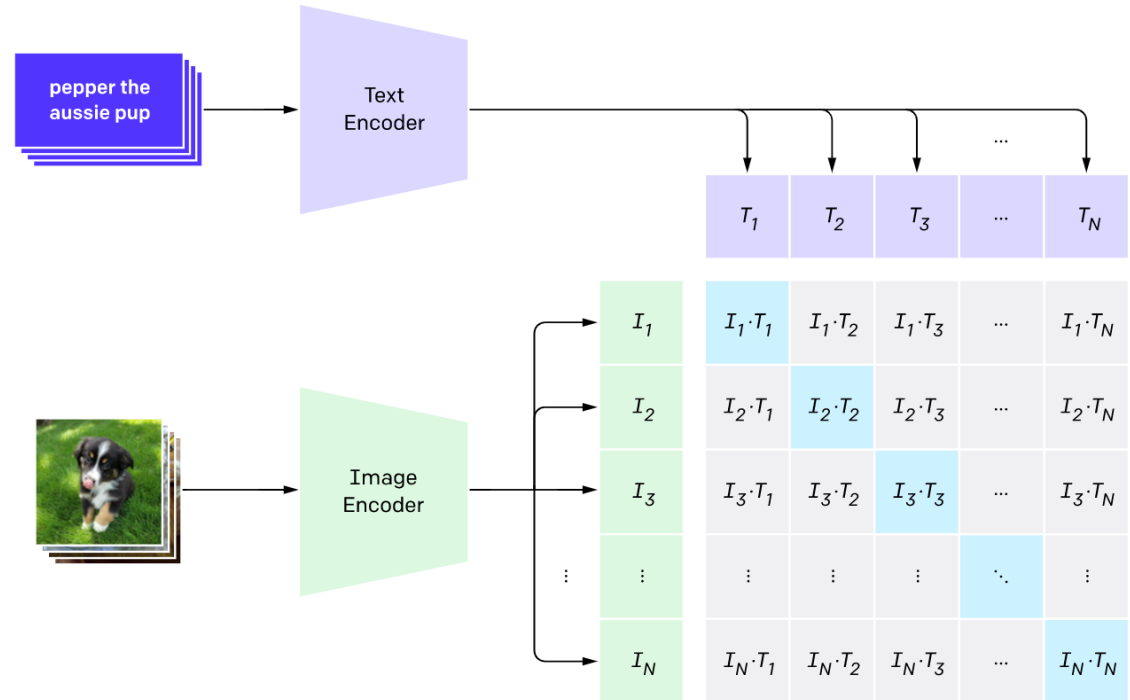
## Word2vec:

Feature space for words



## CLIP:

Common feature space for text and images



# Background: Learn a Feature Space

Why might low dimensional embeddings be useful?

- Example: MNIST digit classification with nearest neighbor

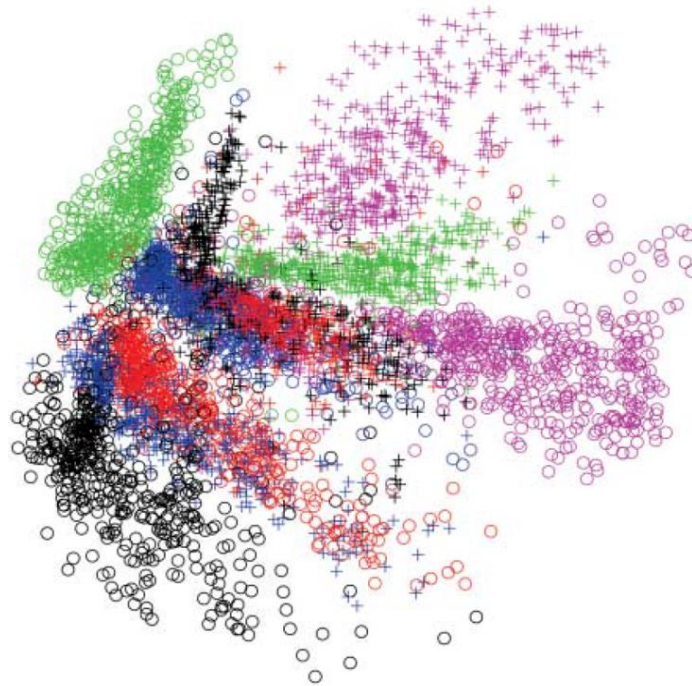
2



# Background: Learn a Feature Space

Why might low dimensional embeddings be useful?

- Example: MNIST digit classification with nearest neighbor





# Background: Measure of Similarity

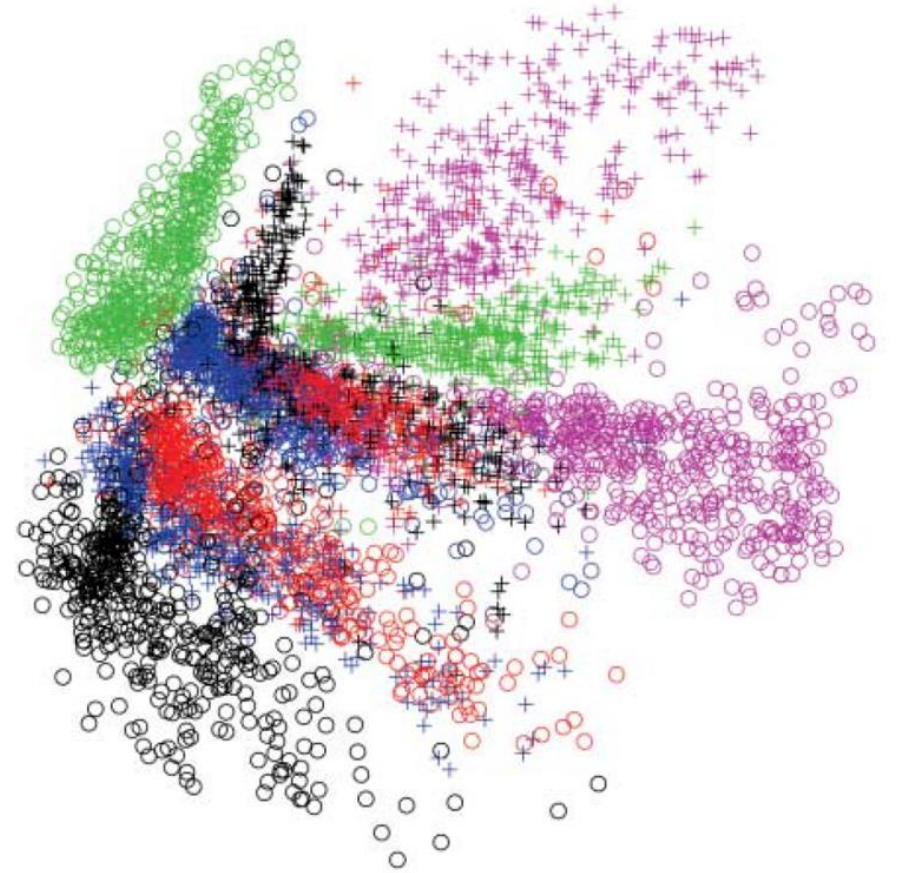
We've been using Euclidean distance

- $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|_2$

Cosine similarity

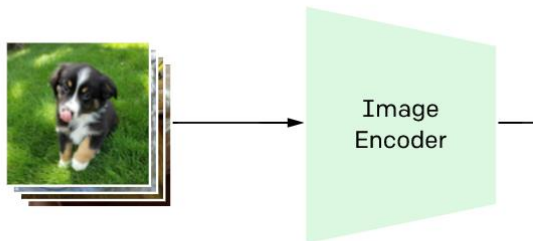
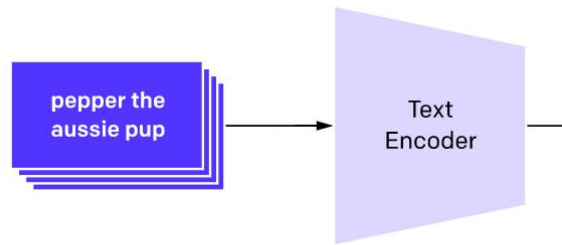
- Two vectors are similar if the angle between them is small

- $d(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T \mathbf{v}$



# Background: Learn a Feature Space

CLIP: Common feature space for text and images

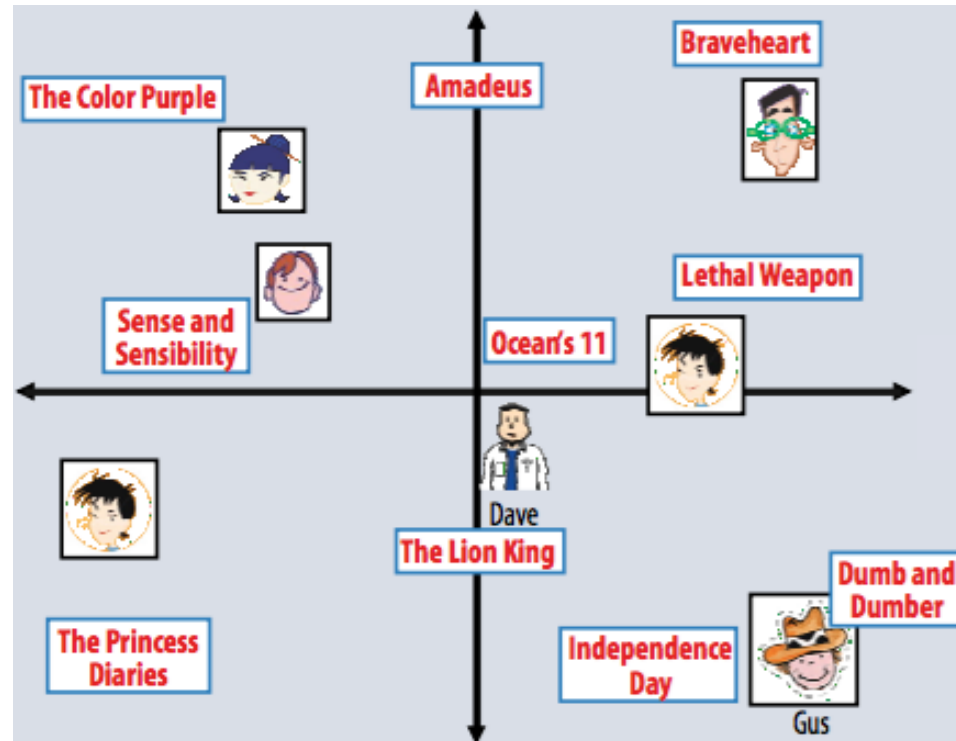


# Recommender System: Matrix Factorization

Learning to map items and users to the same feature space

# Recommender System: Matrix Factorization

Learning to map items and users to the same lower dimensional space



# Recommender System: Matrix Factorization

Optimization: Objective function using only the labels we have

# Poll 1

Is the following optimization a quadratic optimization?

$$\min_{U, V} \sum_{i, j \in \mathcal{S}} \left( r_{ij} - \mathbf{u}^{(i)T} \mathbf{v}^{(j)} \right)^2$$

A.

B.

C.

D.

# Matrix Factorization

Method of *alternating minimization*

$$\min_{U, V} J(U, V) \quad J(U, V) = \sum_{i, j \in \mathcal{S}} \left( r_{ij} - \mathbf{u}^{(i)T} \mathbf{v}^{(j)} \right)^2$$

# Matrix Factorization

Method of *alternating minimization*

$$\min_{U, V} J(U, V) \quad J(U, V) = \sum_{i, j \in \mathcal{S}} (r_{ij} - \mathbf{u}_i^T \mathbf{v}_j)^2$$

- 1)  $\operatorname{argmin}_U J(U, V)$
- 2)  $\operatorname{argmin}_V J(U, V)$



# Matrix Factorization

Method of *alternating minimization*

$$\min_{U,V} J(U, V) \quad J(U, V) = \sum_{i,j \in \mathcal{S}} (r_{ij} - \mathbf{u}_i^T \mathbf{v}_j)^2$$

1)  $\operatorname{argmin}_U J(U, V)$

2)  $\operatorname{argmin}_V J(U, V)$

1) For all  $i$ ,  $\operatorname{argmin}_{\mathbf{u}_i} J(U, V)$

2) For all  $j$ ,  $\operatorname{argmin}_{\mathbf{v}_j} J(U, V)$

# Matrix Factorization

Method of *alternating minimization*

$$\min_{U, V} J(U, V) \quad J(U, V) = \sum_{i, j \in \mathcal{S}} (r_{ij} - \mathbf{u}_i^T \mathbf{v}_j)^2$$

1) For all  $i$ ,  $\operatorname{argmin}_{\mathbf{u}_i} J(U, V)$

$$\mathbf{u}_i^{(t+1)} = (A^T A)^{-1} A^T \mathbf{b}$$

2) For all  $j$ ,  $\operatorname{argmin}_{\mathbf{v}_j} J(U, V)$

$$\mathbf{v}_j^{(t+1)} = (C^T C)^{-1} C^T \mathbf{d}$$

# Matrix Factorization

Method of *alternating minimization*

$$\min_{U, V} J(U, V) \quad J(U, V) = \sum_{i, j \in \mathcal{S}} (r_{ij} - \mathbf{u}_i^T \mathbf{v}_j)^2$$

- 1) For all  $i$ ,  $\operatorname{argmin}_{\mathbf{u}_i} J(U, V) \quad \mathbf{u}_i^{(t+1)} = (A^T A)^{-1} A^T \mathbf{b}$
- 2) For all  $j$ ,  $\operatorname{argmin}_{\mathbf{v}_j} J(U, V) \quad \mathbf{v}_j^{(t+1)} = (C^T C)^{-1} C^T \mathbf{d}$

# Matrix Factorization

First solve:

$$\min_{U, V} J(U, V) \quad J(U, V) = \sum_{i, j \in \mathcal{S}} \left( r_{ij} - \mathbf{u}^{(i)T} \mathbf{v}^{(j)} \right)^2$$

What then?

	Doctor Strange	Star Trek: Beyond	Zootopia
Alita	1		5
BB-8	3	4	
C-3P0	3	5	2

# Matrix Factorization

First solve:

$$\min_{U, V} J(U, V) \quad J(U, V) = \sum_{i, j \in \mathcal{S}} \left( r_{ij} - \mathbf{u}^{(i)T} \mathbf{v}^{(j)} \right)^2$$

What then?

			■		
	■				
			■		
■		■			■
				■	
			■		■
	■				
	■		■		
■					■

# Recommender System: Matrix Factorization

Why is it called matrix factorization?

$$\min_{U, V} \sum_{i, j \in \mathcal{S}} \left( r_{ij} - \mathbf{u}^{(i)T} \mathbf{v}^{(j)} \right)^2$$

# Recommender System: Matrix Factorization

Why is it called matrix factorization?

$$\min_{U,V} \|R - UV^T\|_F^2$$

Sparse labels 😞

	Doctor Strange	Star Trek: Beyond	Zootopia
Alita	1		5
BB-8	3	4	
C-3P0	3	5	2

# Matrix Factorization

Add regularization to avoid overfitting

$$\min_{U, V} J(U, V)$$

$$J(U, V) = \sum_{i, j \in \mathcal{S}} \left( r_{ij} - \mathbf{u}^{(i)T} \mathbf{v}^{(j)} \right)^2$$



# Bias in Recommender Systems

What high-level problems can occur from recommender systems?

# Summary

Recommender systems solve many **real-world** (\*large-scale) **problems**

Collaborative filtering by matrix factorization (MF) is an **efficient** and **effective** approach

(Sparse matrix makes MF more challenging)

MF is just another example of a **common recipe**:

1. define a model
2. define an objective function
3. optimize

## Optimization

- Use alternating minimization
- Add regularization to avoid overfitting