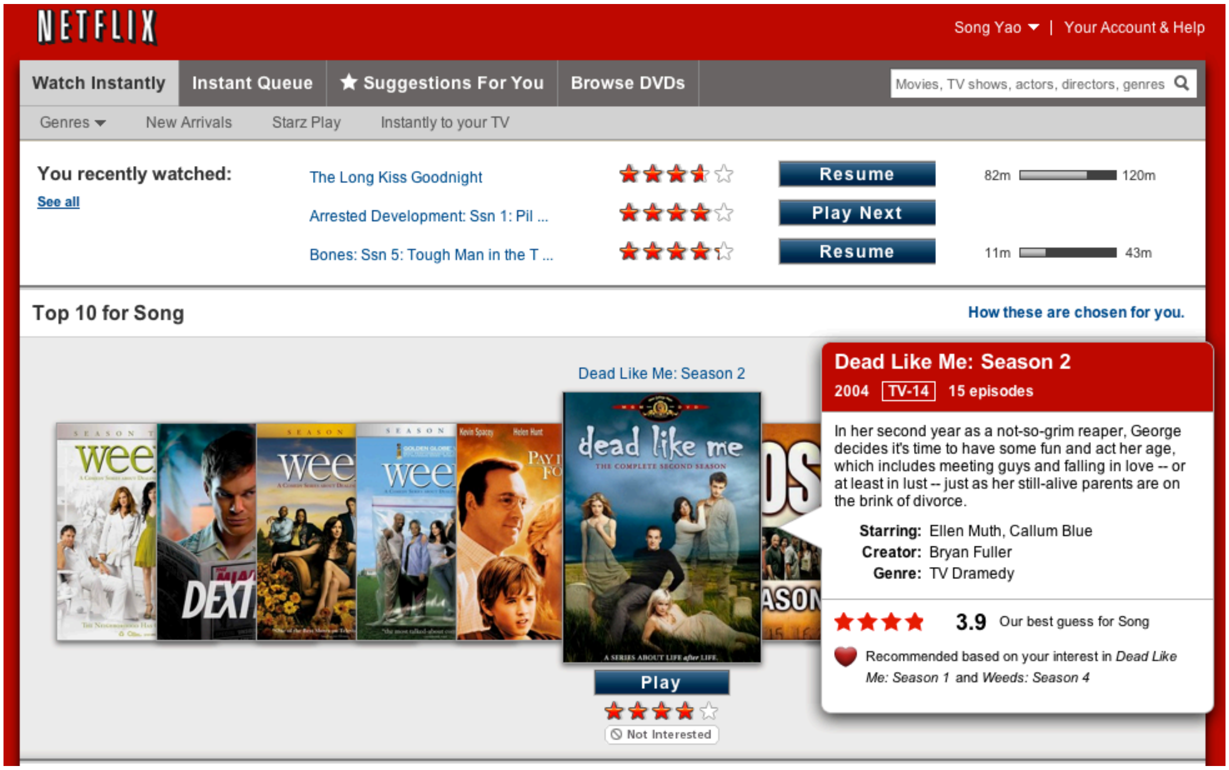
Recommendation Systems: Learning from Ratings and More

Professor Song Yao
Olin Business School

Customer Analytics



You recently watched:
See all

The Long Kiss Goodnight	82m	120m
Arrested Development: Ssn 1: Pil ...		
Bones: Ssn 5: Tough Man in the T ...	11m	43m

Top 10 for Song

Dead Like Me: Season 2

Dead Like Me: Season 2
2004 TV-14 15 episodes

In her second year as a not-so-grim reaper, George decides it's time to have some fun and act her age, which includes meeting guys and falling in love -- or at least in lust -- just as her still-alive parents are on the brink of divorce.

Starring: Ellen Muth, Callum Blue
Creator: Bryan Fuller
Genre: TV Dramedy

3.9 Our best guess for Song

Recommended based on your interest in *Dead Like Me: Season 1* and *Weeds: Season 4*

Ratings data give more detailed insight into consumer preferences than simple 0/1 purchase decisions

STRUCTURE OF RATINGS DATA

	Products										
Users	A	B	C	D	E	F	G	H	I	J	...
1	4	1	3	5	1	5	1	1	3	2	
2	1	3	1	4	5	2	3	1	2	4	
3	3	2	5	4	1	5	2	4	4	3	
4	1	4	2	4	1	4	4	3	5	2	
5	3	1	2	5	4	5	2	1	1	3	
6	4	3	1	5	1	5	4	1	3	2	
7	1	4	1	5	4	5	5	3	3	2	
8	2	3	5	3	1	4	2	5	3	4	
9	1	3	1	4	4	2	3	2	2	5	
10	1	3	2	2	4	3	3	4	1	5	
...											

If ratings is all we observe, the only way to predict ratings of a consumers is to use ratings of other consumers

RATINGS PREDICTION

	Products									
Users	A	B	C	D	E	F	G	H	I	J
1	4	1	3	5	1	5	1	1	3	2
2	1	3	1	4	5	2	3	1	2	4
3	3	2	5	4	1	5	2	4	4	3
4	1	4	2	4	1	4	4	3	5	2
5	3	1	2	5	4	5	2	1	1	3
6	4	3	1	5	1	5	4	1	3	2
7	1	4	1	5	4	5	5	3	3	2
8	2	3	5	3	1	4	2	5	3	4
9	1	3	1	4	4	2	3	2	2	5
10	1	3	2	2	4	3	3	4	1	5
11	.	.	.	.	.	.	.	.	.	.

The simplest approach is to average over other users' ratings

RATINGS PREDICTION

	Products									
Users	A	B	C	D	E	F	G	H	I	J
1	4	1	3	5	1	5	1	1	3	2
2	1	3	1	4	5	2	3	1	2	4
3	3	2	5	4	1	5	2	4	4	3
4	1	4	2	4	1	4	4	3	5	2
5	3	1	2	5	4	5	2	1	1	3
6	4	3	1	5	1	5	4	1	3	2
7	1	4	1	5	4	5	5	3	3	2
8	2	3	5	3	1	4	2	5	3	4
9	1	3	1	4	4	2	3	2	2	5
10	1	3	2	2	4	3	3	4	1	5
11	2.1	2.7	2.3	4.1	2.6	4	2.9	2.5	2.7	3.2

- Average over other user's rating for the same item $r_t(p) = \frac{\sum_{u \neq t} r_u(p)}{n}$
  - t = target user
  - p = product
  - u = other users
  - n = number of items

In collaborative filtering we use ratings from other people, but more so from users who are similar to the target user

COLLABORATIVE FILTERING IDEA

- Recognize that **consumers have different preferences**
  - Some prefer action movies, others prefer British period dramas...
  - Some prefer movies with Tom Cruise, others prefer movies with Daniel Craig
  - Some prefer romantic comedies **and** horror movies but **not** action movies
- If we try to predict what movies a target user would like, we should
  - **positively** weight ratings from people who a similar
  - **negatively** weight ratings from people who have "inverse" preferences
  - **not weight** ratings from people whose ratings seem to be unrelated

To use collaborative filtering we need to observe <u>some</u> ratings for the target consumer

REQUIRED RATINGS DATA

	Items									
Users	A	B	C	D	E	F	G	H	I	J
1	4	1	3	5	1	5	1	1	3	2
2	1	3	1	4	5	2	3	1	2	4
3	3	2	5	4	1	5	2	4	4	3
4	1	4	2	4	1	4	4	3	5	2
5	3	1	2	5	4	5	2	1	1	3
6	4	3	1	5	1	5	4	1	3	2
7	1	4	1	5	4	5	5	3	3	2
8	2	3	5	3	1	4	2	5	3	4
9	1	3	1	4	4	2	3	2	2	5
10	1	3	2	2	4	3	3	4	1	5
11	2	3	2	5	1	.	.	.	.	.

We need to find ways to give different users different weights in the calculations

WEIGHTS IN COLLABORATIVE FILTERING

- Measure whether two consumers have similar preferences:
  - For products that **both consumers** rated, are consumers' ratings correlated?
    $\mathrm{corr}(r_t, r_u)$ can be >0, <0, or 0; **for now**, we assume it is >=0 as is typical
- Create a "user weight" to weight the contribution of this user's rating in forming the recommendation

$$w(t,u) = \frac{\mathrm{corr}(r_t, r_u)}{\sum_{u \neq t} \mathrm{corr}(r_t, r_u)}$$

- ratings correlation between target user "t" and user "u"
- sum of all other ratings correlations

EXAMPLE

- target user and 3 other users

$\mathrm{corr}(r_t, r_1) = .05$

$\mathrm{corr}(r_t, r_2) = .8$

$\mathrm{corr}(r_t, r_3) = .3$

$$w(t,1) = \frac{0.05}{0.05 + 0.8 + 0.3} = 0.04$$

$$w(t,2) = \frac{0.8}{1.15} = 0.70$$

$$w(t,3) = \frac{0.3}{1.15} = 0.26$$

We can use the calculated weights to combine the ratings of other users to form a recommendation for the target user

COMBINING RATINGS: SIMPLE APPROACH

- Predicted rating of the target user for product *p* is weighted sum of ratings of all other users for product *p*

$$r_t(p) = \sum_{u \neq t} r_u(p) * w(t,u)$$

EXAMPLE

- target user and 3 other users

$$\begin{array}{lll}
\mathrm{corr}(r_t, r_1) = .05 & \mathrm{corr}(r_t, r_2) = .8 & \mathrm{corr}(r_t, r_3) = .3 \\
w(t,1) = \frac{0.05}{1.15} = 0.04 & w(t,2) = \frac{0.8}{1.15} = 0.70 & w(t,3) = \frac{0.3}{1.15} = 0.26 \\
r_1(p) = 1 & r_2(p) = 5 & r_3(p) = 3
\end{array}$$

$$r_t(p) = 0.04 * 1 + 0.7 * 5 + 0.26 * 3 = \boxed{4.32}$$

Because of weights, collaborative filtering can generate very different prediction for different users from the same ratings

EXAMPLE

- target user and 3 other users

$$\begin{array}{lll}
\mathrm{corr}(r_t, r_1) = .05 & \mathrm{corr}(r_t, r_2) = .8 & \mathrm{corr}(r_t, r_3) = .3 \\
w(t,1) = \frac{0.05}{1.15} = 0.04 & w(t,2) = \frac{0.8}{1.15} = 0.70 & w(t,3) = \frac{0.3}{1.15} = 0.26 \\
r_1(p) = 1 & r_2(p) = 5 & r_3(p) = 3
\end{array}$$

$$r_t(p) = 0.04 * 1 + 0.7 * 5 + 0.26 * 3 = \boxed{4.32}$$

- **different target user** and **the same** 3 other users

$$\begin{array}{lll}
\mathrm{corr}(r_t, r_1) = .65 & \mathrm{corr}(r_t, r_2) = .2 & \mathrm{corr}(r_t, r_3) = .03 \\
w(t,1) = \frac{0.65}{0.88} = 0.74 & w(t,2) = \frac{0.2}{0.88} = 0.23 & w(t,3) = \frac{0.03}{0.88} = 0.03 \\
r_1(p) = 1 & r_2(p) = 5 & r_3(p) = 3
\end{array}$$

$$r_t(p) = 0.74 * 1 + 0.23 * 5 + 0.03 * 3 = \boxed{1.98}$$

What if consumers use rating scales differently?

- **Problem:** target users rates from 2-4, other users from 2-5
- **Solution:** Instead of predicting absolute ratings, predict *deviations from the target user's mean ratings, adjusted by standard deviation (z-score)*

Average rating for target user

Standard deviation of ratings for user u

$$\frac{r_t(p) - \bar{r}_t}{sd(r_t)} = \sum_{u \neq t} \frac{r_u(p) - \bar{r}_u}{sd(r_u)} * w(t,u)$$

Average rating for user u

Standard deviation of ratings for user u

EXAMPLE

$\text{corr}(r_t, r_1) = .05$

$w(t,1) = \frac{0.05}{1.15} = 0.04$

$r_1(p) = 1$

$\bar{r}_1 = 2.3 \quad sd(r_1) = 0.85$

$\text{corr}(r_t, r_2) = .8$

$w(t,2) = \frac{0.8}{1.15} = 0.70$

$r_2(p) = 5$

$\bar{r}_2 = 3.3 \quad sd(r_2) = 1.25$

$\text{corr}(r_t, r_3) = .3$

$w(t,3) = \frac{0.3}{1.15} = 0.26$

$r_3(p) = 3$

$\bar{r}_3 = 4.1 \quad sd(r_3) = 0.7$

$$\frac{r_t(p) - \bar{r}_t}{sd(r_t)} = 0.04 * \frac{(1-2.3)}{0.85} + 0.7 * \frac{(5-3.3)}{1.25} + 0.26 * \frac{(3-4.1)}{0.7} = \boxed{0.48}$$

$$\bar{r}_t = 2.9, sd(r_t) = 1.1 \Rightarrow \frac{r_t(p) - 2.9}{1.1} = 0.48 \Leftrightarrow \boxed{r_t(p) = 3.43}$$ (was 4.32)

What if consumers ratings are negatively correlated?

TWO APPROACHES FOR NEGATIVELY CORRELATED RATINGS

1. Ignore users with negatively correlated ratings: For a given target user, only consider other users with ratings that are positively correlated with that target user.

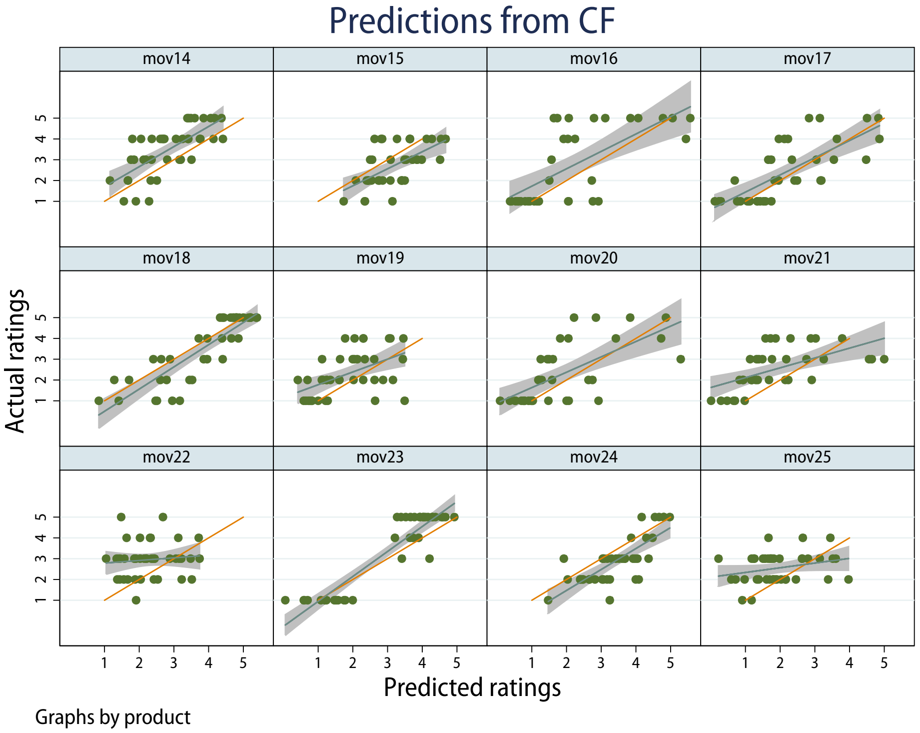
2. If the ratings scale is centered around zero, use a modified weight function

$$w(t,u) = \frac{\text{corr}(r_t, r_u)}{\sum_{u \neq t} |\text{corr}(r_t, r_u)|}$$

and apply prior combination formulae.

- Ratings on a {1,...,*k*} scale are not centered around zero but can easily be shifted by subtracting (*k*+1)/2 so that they are (e.g., {1,2,3,4,5} --> {-2,-1,0,1,2}).
- Standardized (*z*-score) ratings are centered around zero.

Collaborative filtering has predictive power, but predictions are not always correct



There are a variety of challenges in collaborative filtering

ISSUES IN COLLABORATIVE FILTERING

- **How does one get the ratings?**
  - Input explicitly by the user; facilitation by user interface is key
  - Implicit ratings from observations of clickstream behavior
- **Ratings data can be corrupted**
  - ratings from others (shared accounts)
  - purchases for others (when using clickstream behavior)
- **How persistent is the person-to-person similarity across product categories or genres?**
  - Netflix first separates films into several clusters
  - Collaborative filtering is used to recommend films from each cluster, using ratings and similarity data for films in that cluster only

There are many variants in collaborative filtering

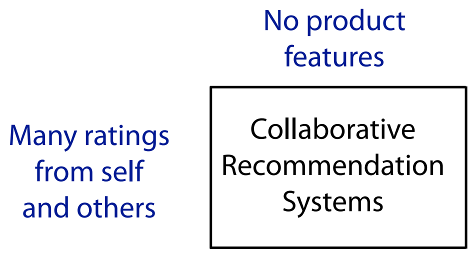
VARIANTS IN COLLABORATIVE FILTERING

- **Supplementing ratings data**
  - So far similarity between users established only based on ratings
  - When ratings overlap is limited, one can supplement user similarity calculations with other characteristics (e.g. demographics)
- **User-based vs. item-based collaborative filtering**
  - So far:
    - Calculate similarity between users based on ratings
    - Use *other items'* ratings by *similar users* to form prediction
  - Alternative (if few common ratings between users):
    - Calculate similarity between items based on ratings of others
    - Use *own* ratings of *similar items* to form prediction

---

While CF requires no knowledge of products, it requires other's ratings on many relevant products

APPROACHES TO RATING PREDICTIONS



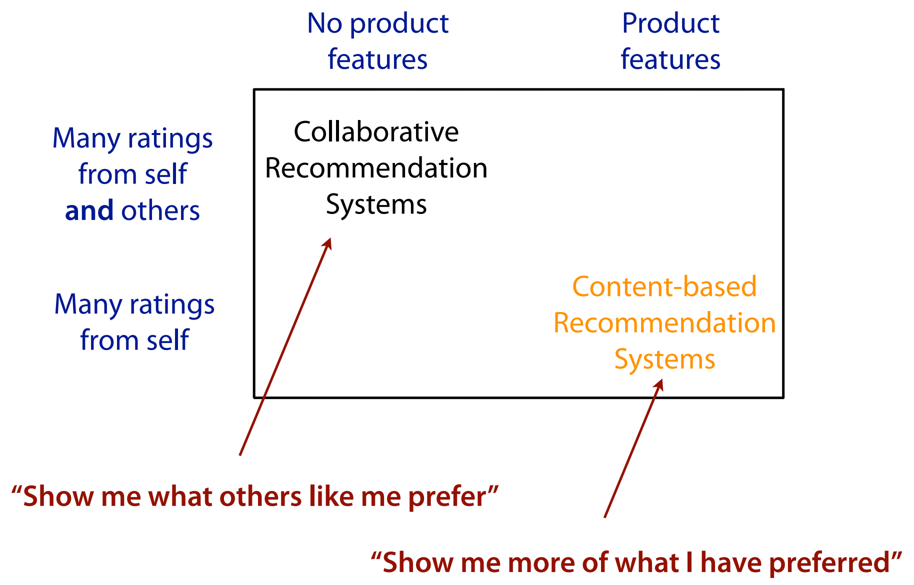
No product features

Many ratings from self and others

Collaborative Recommendation Systems

What if we don’t have relevant ratings from others?

APPROACHES TO RATING PREDICTIONS



No product features

Product features

Many ratings from self **and** others

Many ratings from self

Collaborative Recommendation Systems

Content-based Recommendation Systems

“Show me what others like me prefer”

“Show me more of what I have preferred”

When we have many product attributes/features and ratings by the user we no longer need the ratings of others

DATA OF A SINGLE CUSTOMER

		Attributes/Features				
Products	Ratings	1	2	3	4	5
1	3	1	7	9	9	6
2	5	9	5	6	6	5.1
3	1	1	4	10	7	2
4	5	10	6	1	2	1.3
5	2	3	7	5	5	4
6	3	7	9	8	7	2.4
7	4	9	8	8	4	3.6
8	1	2	10	10	9	9.3
9	4	6	1	7	4	3
10	2	5	7	9	10	1.1
11	?	2	3	4	1	3.1

When we have many product attributes/features and ratings by the user we no longer need the ratings of others

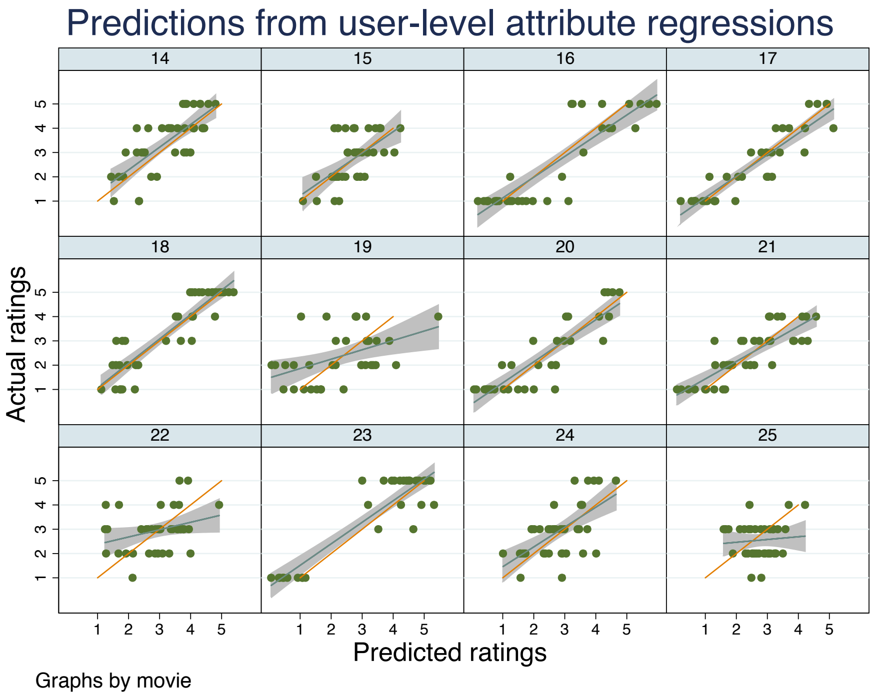
DATA OF A SINGLE CUSTOMER

		Attributes/Features				
Products	Ratings	1	2	3	4	5
1	3	1	7	9	9	6
2	5	9	5	6	6	5.1
3	1	1	4	10	7	2
4	5	10	6	1	2	1.3
5	2	3	7	5	5	4
6	3	7	9	8	7	2.4
7	4	9	8	8	4	3.6
8	1	2	10	10	9	9.3
9	4	6	1	7	4	3
10	2	5	7	9	10	1.1
11	?	2	3	4	1	3.1

Example of content-based recommendation system: *User-level regression*

$$\text{Rating}=a+b\bullet\text{Attr}_1+c\bullet\text{Attr}_2+d\bullet\text{Attr}_3+e\bullet\text{Attr}_4+f\bullet\text{Attr}_5$$

Content-based recommendation systems also have predictive power



The strength of collaborative vs. content-based recommendation systems depends on the strength of data

DATA VS. PERFORMANCE

- Low ratings overlap b/w focal customer and others
  --> **Collaborative recommendation** performs poorly
    - Correlations estimates are noisy
- Few or badly measured attributes
  --> **Content-based recommendation** perform poorly
    - Too little of the rating is explained by independent variables

---

What is we have data on other users <u>and</u> data on product features?

APPROACHES TO RATING PREDICTIONS

	No product features	Product features
Many ratings from self **and** others	Collaborative Recommendation Systems	Hybrid Recommendation Systems
Many ratings from self		Content-based Recommendation Systems

~~Azure ML: “Matchbox Recommender”~~

Demo code inspired by Matchbox Recommener

Can we recommend anything if we only know one rating?

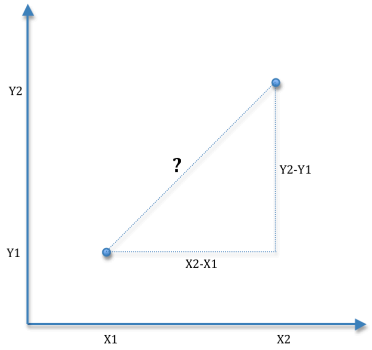
DATA OF A SINGLE CUSTOMER

		Attributes/Features				
Products	Ratings	1	2	3	4	5
1	Like	1	7	9	9	6
2	?	9	5	6	6	5.1
3	?	1	4	10	7	2
4	?	10	6	1	2	1.3
5	?	3	7	5	5	4
6	?	7	9	8	7	2.4
7	?	9	8	8	4	3.6
8	?	2	10	10	9	9.3
9	?	6	1	7	4	3
10	?	5	7	9	10	1.1
11	?	2	3	4	1	3.1

We have to rely solely on the similarity of products

Pandora Music Genome Project (US Patent No. 7,003,515)

- For every other product, calculate a “distance” to this product
- Sort the products based on their distances to this product
- How to compute the distance?
  - Each product is defined by its attributes
  - The difference between the attributes of a pair of products is the “distance”



$$Distance = \sqrt{(X2 - X1)^2 + (Y2 - Y1)^2}$$

Calculating similarity is very easy

EXAMPLE: RECOMMENDING SIMILAR MOVIES

		Attributes								**(Ratings Difference)^2**								
Title	**rating**	dark	violence	crime	mysteri	romance	feelgood	goofy	family	dark	violence	crime	mysteri	romance	feelgood	goofy	family	Distance
Shutter Island	Like	4	4	4	5	1	1	1	1	0	0	0	0	0	0	0	0	0.0
Bourne Identity	?	3	4	3	2	2	2	1	1	1	0	1	9	1	1	0	0	3.6
Seven	?	5	5	5	5	1	1	1	1	1	1	1	0	0	0	0	0	1.7
8MM	?	5	5	5	5	1	1	1	1	1	1	1	0	0	0	0	0	1.7
Heartbreaker	?	1	1	1	1	5	5	4	2	9	9	9	16	16	16	9	1	9.2
When Harry Met Sally	?	1	1	1	1	5	5	3	3	9	9	9	16	16	16	4	4	9.1
Good Will Hunting	?	1	2	1	1	4	5	1	3	9	4	9	16	9	16	0	4	8.2
Dead Like Me	?	2	3	2	2	2	3	3	2	4	1	4	9	1	4	4	1	5.3
Kong Fu Panda	?	1	2	1	1	1	4	4	5	9	4	9	16	0	9	9	16	8.5
Sex and the City	?	1	1	1	1	4	4	3	1	9	9	9	16	9	9	4	0	8.1
North and South	?	1	1	1	1	5	5	1	3	9	9	9	16	16	16	0	4	8.9
Pride and Prejudice	?	1	1	1	1	5	5	1	3	9	9	9	16	16	16	0	4	8.9
Jane Eyre	?	3	1	1	1	5	4	1	2	1	9	9	16	16	9	0	1	7.8
Up	?	1	1	1	1	2	5	3	5	9	9	9	16	1	16	4	16	8.9
Salt	?	3	4	3	4	1	3	1	1	1	0	1	1	0	4	0	0	2.6
Emma	?	1	1	1	1	5	5	1	4	9	9	9	16	16	16	0	9	9.2
Chocolat	?	1	1	1	1	5	5	2	4	9	9	9	16	16	16	1	9	9.2

(C4-C$3)^2

SQRT(SUM(K4:R4))

Pandora's key ingredient is a comprehensive system of attributes to categorize music

PANDORA MAJOR ATTRIBUTES

- 1 Structures/Composition
- 2 Rhythm/Meter
- 3 Ostinato
- 4 Roots
- 5 Tonality
- 6 Instrumentation
- 7 Feel
- 8 Musical qualities
- 9 Leanings/stylings
- 10 Recording techniques
- 11 Influences
- 12 Instrumental Ensembles
  - 12.1 String section
  - 12.2 Brass and/or Horns sections
  - 12.3 Percussion Sections
- 13 Individual Instruments
  - 13.1 Bass Guitar
  - 13.2 Contrabass
  - 13.3 Drums
  - 13.4 Cymbals
  - 13.5 Guitar - Either
  - 13.6 Guitar (Acoustic)
  - 13.7 Guitar (Electric)
  - 13.8 Keyboarded
    - 13.8.1 Accordion
    - 13.8.2 Harpsichord
    - 13.8.3 Organ
    - 13.8.4 Piano
    - 13.8.5 Synth
      - 13.8.5.1 Imitative Synthesis
  - 13.9 Non-Pitched Percussion
    - 13.9.1 Hand Percussion
  - 13.10 Horns
    - 13.10.1 Trombone
    - 13.10.2 Trumpet
  - 13.11 Idiophone
  - 13.12 String
  - 13.13 Woodwind
    - 13.13.1 Sax
      - 13.13.1.1 Soprano
      - 13.13.1.2 Alto
      - 13.13.1.3 Tenor
      - 13.13.1.4 Baritone
  - 13.14 Other
- 14 Lyrical content
- 15 Vocals
  - 15.1 Male
  - 15.2 Female

Pandora's key ingredient is a comprehensive system of attributes to categorize music

PANDORA MINOR ATTRIBUTES

Structures/Composition

- Basic Rock Song Structures
- Big Band Arrangements
- Buildup/Breakdown
- Chromatic Harmonic Structure
- Compelling Intensity
- Epic Buildup/Breakdown
- Great Musicianship
- Groove Based Composition
- Interesting Song Structure
- Intricate Arranging
- Lead Big Band
- Melodic Songwriting
- Minimalist Arrangements
- Orchestral Arranging
- Repetitive Song Structure
- Strongly Dramatic Aesthetic
- Subtle Buildup/Breakdown

Rhythm/Meter

- Danceable Grooves
- Four-Four Time Signature
- Hard Swingin' Rhythm
- Heavy Syncopation
- Intricate Rhythms
- Meter Complexity
- Mild Rhythmic Syncopation
- Triple Meter Style
- Triple Note Feel
- Twelve-Eight Time Signature
- Unsyncopated Ensemble Rhythms
- Use of Groove
- Varying Tempo and Time Signatures
- Rhythmic Intro

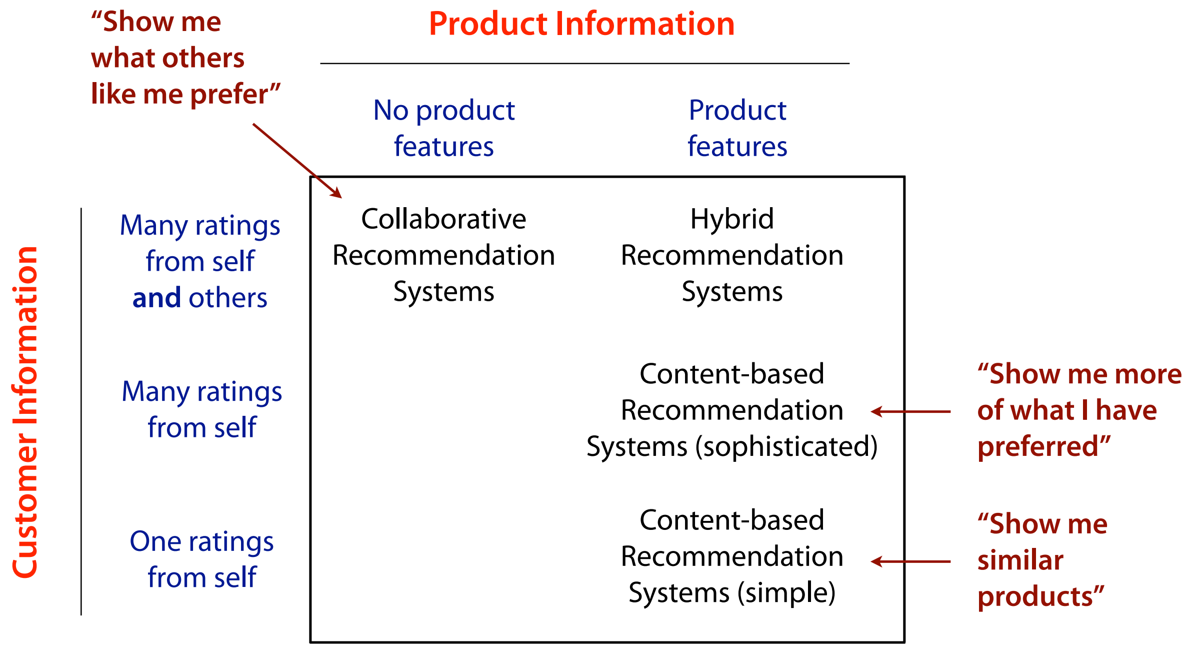
Ostinato

- Acoustic Sonority
- Acousti-synthetic Sonority
- Chordal Patterning
- Dominant Use of Riffs
- Electro-Synthetic Sonority
- Extensive Vamping
- Highly Synthetic Sonority
- Intricate Melodic Phrasing
- Knack for Catchy Hooks
- Melodic Songwriting
- Modal Harmony
- Overall Meditative Sound
- Repetitive Chorus
- Repetitive Melodic Phrasing
- Repetitive Verse
- Synth-acoustic Sonority
- Synth-electric Sonority
- Tonal Harmony

Roots

- Acid Jazz Roots
- Afro-Latin Roots
- Basic Rap Roots
- Blues Roots
- Caribbean Roots
- Classic Jazz Roots
- Club Rap Roots
- Cool Jazz Roots
- Country Roots
- Disco Roots
- East Coast Rap Roots
- Electronica Roots
- Folk Roots
- Funk Roots
- Hard Bop Roots
- Hard Rock Roots
- Hawaiian Roots
- House Roots
- Industrial Roots
- Meso-American Roots
- Midwest Rap Roots
- New Orleans Jazz Roots
- Old School Roots
- Punk Roots
- Rock & Roll Roots
- Ska Roots
- Southern Rap Roots
- Swing Era Roots
- Techno Roots
- Trance Roots
- Trip Hop Roots
- West Coast Rap Roots
- ...
- ...

You need to know your data to know how to predict ratings



"Show me what others like me prefer"

		Product Information	
		No product features	Product features
Customer Information	Many ratings from self **and** others	Collaborative Recommendation Systems	Hybrid Recommendation Systems
	Many ratings from self		Content-based Recommendation Systems (sophisticated)
	One ratings from self		Content-based Recommendation Systems (simple)

"Show me more of what I have preferred"

"Show me similar products"