Week 2: Basic Concepts and Similarity Measures

CS3448: Recommender Systems /
CSX4207/ITX4207: Decision Support and Recommender Systems

Asst. Prof. Dr. Rachsuda Setthawong

Asst. Prof. Dr. Rachsuda Setthawong

Objectives

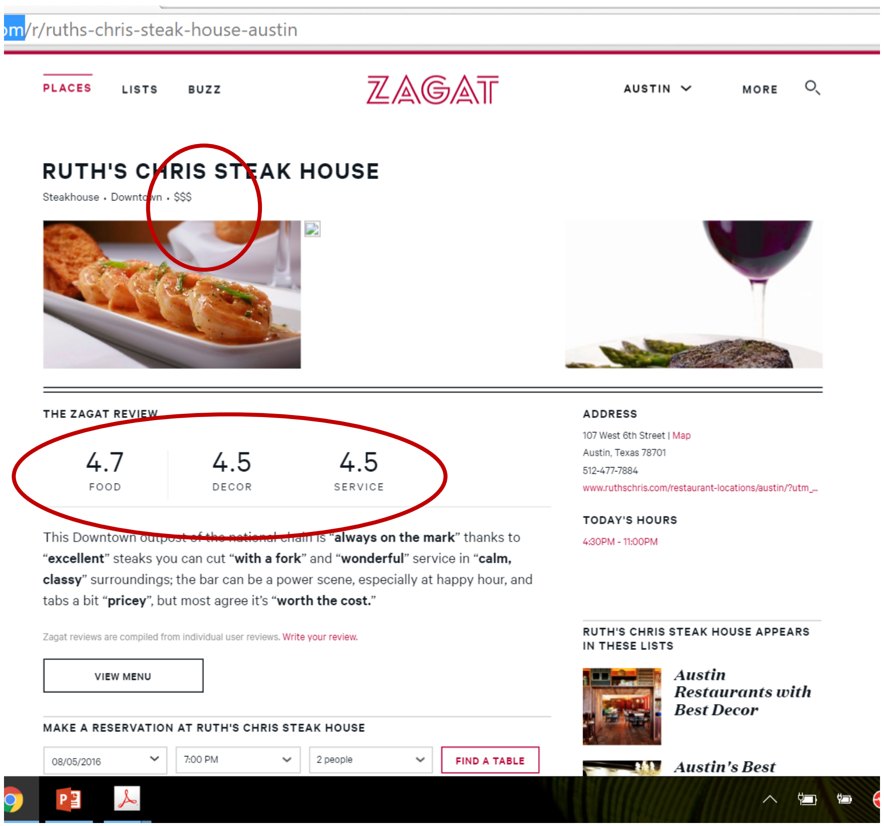
- To determine weak points of non-personalized recommendations
- To understand types of data used in generating recommendations in personalized RSs
- To understand basic concepts used in generating recommendations in personalized RSs
- To understand scoring and ranking of data
- To understand basic similarity measures

Outlines

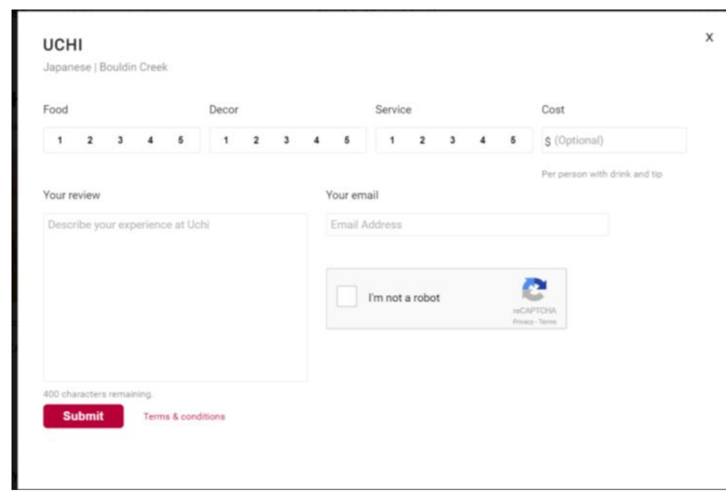
- Weak Points of Non-Personalized RS
- Preferences and Ratings
- Predictions and Recommendations
- Scoring and Ranking
- Basic Similarity Measures

Asst. Prof. Dr. Rachsuda Setthawong

Revisiting Non-Personalized RS



- Generating Recommendations
  - Rating = {1, 2, 3, 4, 5}
  - Score = MEAN(ratings)



Asst. Prof. Dr. Rachsuda Setthawong

Weak Points of Non-Personalized RS

- Side-effects of averaging
  - Moderate restaurants with good scores (many opinions are exaggerated.)
  - Great restaurants with moderate score (many typical people may not like niche recipe.)
  - Not aware of concept drift (Some restaurant was very good in the past, but it is worse now.)
- Recommendations not customized to individual needs
  - E.g.,
    - Suggest top 10 greatest hits albums all in Pop to A Cappella's fan
    - Suggest products in supermarkets based on best selling ones to every customer

Outlines

- Weak Points of Non-Personalized RS
- Preferences and Ratings
- Predictions and Recommendations
- Scoring and Ranking
- Basic Similarity Measures

Personalized RSs

- Make use of individuals' information to generate recommendations
  - Preferences
  - Ratings

Asst. Prof. Dr. Rachsuda Setthawong

Types of Preferences

Explicit	Implicit
• Rating	• Click to view
• Vote	• Buy
• Review	• Follow

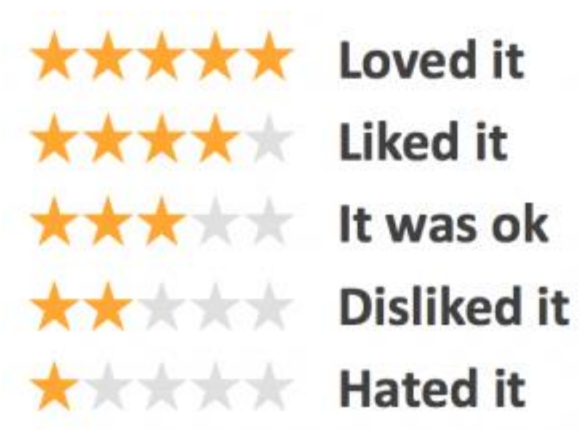
How to classify user preferences?
How do you classify
- 'post', 'like', 'comment' and 'share' in Facebook?
- 'tweet' and 'retweet' in Twitter?

Explicit Ratings

- Ask users straightforwardly for scoring a given item.
- Examples of ratings
  - **Star ratings**
    - Typical 5 stars (with or with out half star)
  - **Likert scale**
    - A typical five-level Likert item: 'Strongly disagree', 'Disagree', 'Neither agree nor disagree', 'Agree', 'Strongly agree'
    - Different level Likert scales[1]: 3, 7, 10, etc.

[1]More info: http://www.socialresearchmethods.net/kb/scallik.php

Examples of Star Rating



Examples of Likert Scales

Likert Scales

Please fill in the number that represents how you feel about the computer software you have been using

I am satisfied with it

1	2	3	4	5
Strongly Agree	Agree	Neither	Disagree	Strongly Disagree

It is simple to use

1	2	3	4	5
Strongly Agree	Agree	Neither	Disagree	Strongly Disagree

It is fun to use

1	2	3	4	5
Strongly Agree	Agree	Neither	Disagree	Strongly Disagree

It does everything I would expect it to do

1	2	3	4	5
Strongly Agree	Agree	Neither	Disagree	Strongly Disagree

I don't notice any inconsistencies as I use it

1	2	3	4	5
Strongly Agree	Agree	Neither	Disagree	Strongly Disagree

Likert Scales

Please circle the number that represents how you feel about the computer software you have been using

I am satisfied with it
Strongly Disagree ---1---2---3---4---5---6---7--- Strongly Agree

It is simple to use
Strongly Disagree ---1---2---3---4---5---6---7--- Strongly Agree

It is fun to use
Strongly Disagree ---1---2---3---4---5---6---7--- Strongly Agree

It does everything I would expect it to do
Strongly Disagree ---1---2---3---4---5---6---7--- Strongly Agree

I don't notice any inconsistencies as I use it
Strongly Disagree ---1---2---3---4---5---6---7--- Strongly Agree

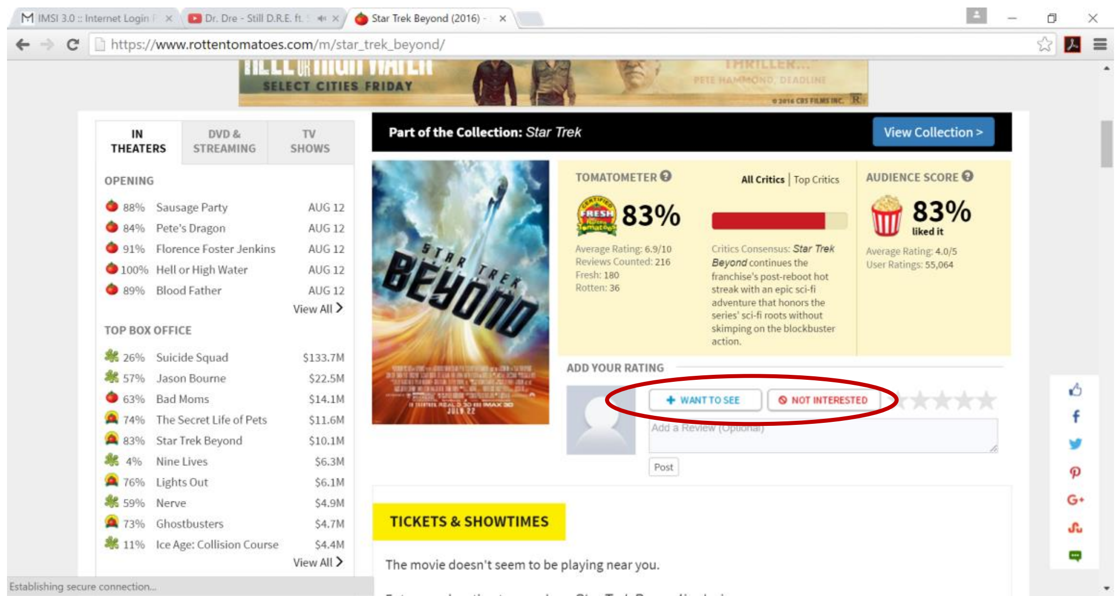
It is very user friendly
Strongly Disagree ---1---2---3---4---5---6---7--- Strongly Agree

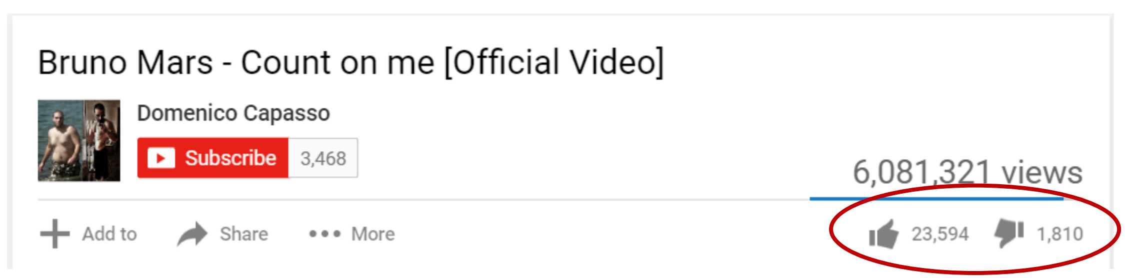
Sources: http://www.hkadesigns.co.uk/websites/msc/reme/likert.htm

Asst. Prof. Dr. Rachsuda Setthawong

Vote

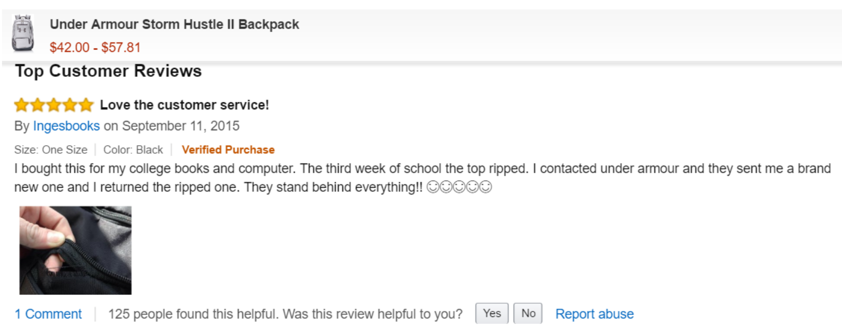
- Likes
- Thumbs





Sources: https://www.rottentomatoes.com
https://www.youtube.com

Review



Sources: https://www.amazon.com
https://www.hotels.com

Asst. Prof. Dr. Rachsuda Setthawong

Timing to Give Ratings

expectation

Before experiencing the item

consumption

During or immediately after experiencing the item

memory

Some time after experience

Issues of Rating Usages

- Reliability and accuracy
- User preferences drifting
- Rating's meaning

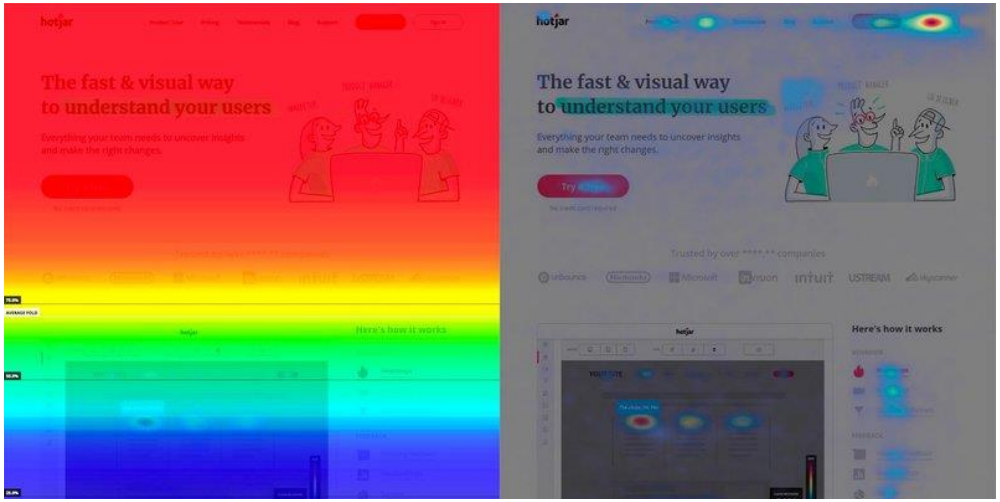
Source: https://xkcd.com/1098/

Asst. Prof. Dr. Rachsuda Setthawong

Implicit Preferences

- Observe from user behaviors
- Not easily interpreted as ratings
- Give more details beside ratings
  - E.g.,
    - Reading/watching time
    - Click on link/ad
    - Add to cart/buy
    - Search/share content

Heatmaps



Source: https://www.hotjar.com/behavior-analytics-software/
How tracking user behavior on your website can improve customer experience: https://www.hotjar.com/blog/user-behavior/

Asst. Prof. Dr. Rachsuda Setthawong

Advantages and Disadvantages

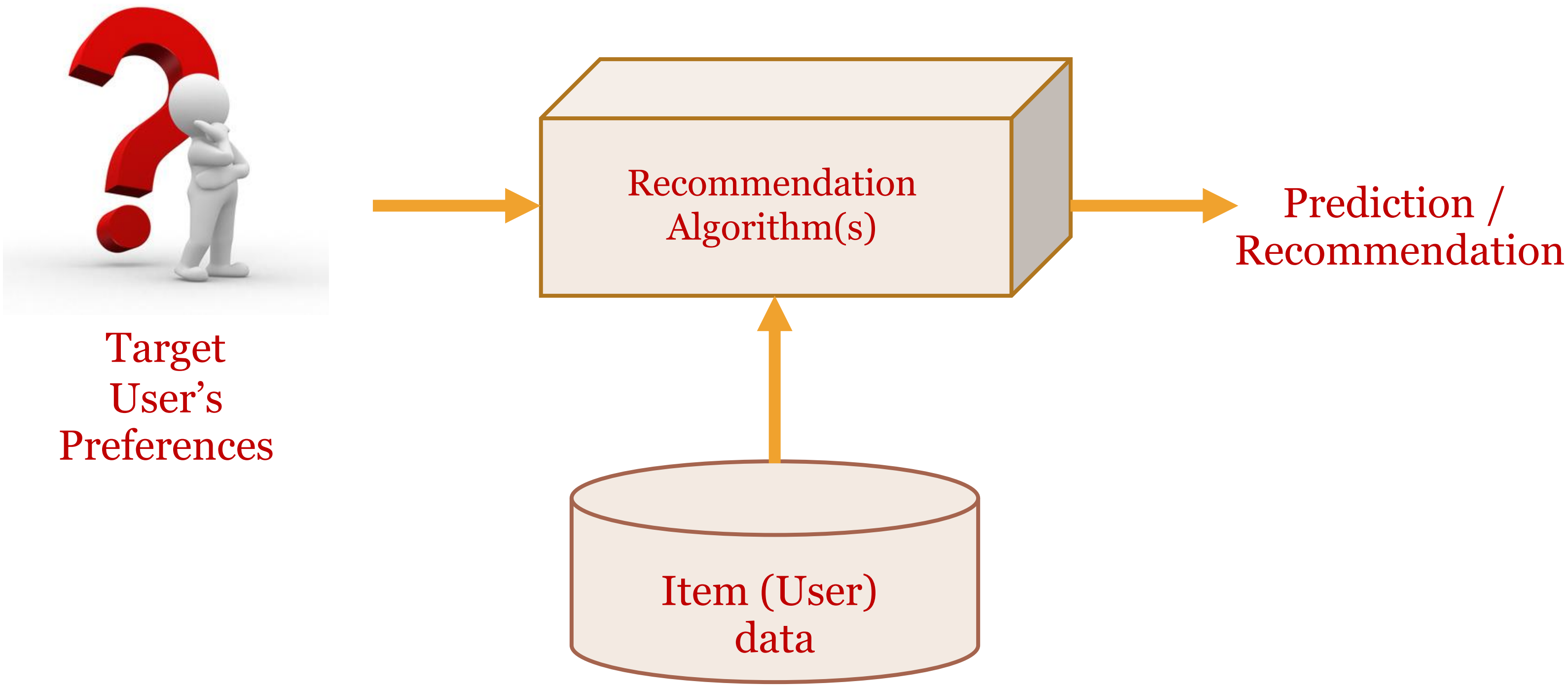
	Explicit Ratings	Implicit Preference
Used to generate personalized recommendations	Yes	Yes
Easy to interpretation	Yes	No
Effort to collect data	require user efforts	great amount of data available
Reliability and accuracy	more	less
User preferences drifting	may be concerned	-
Scale/represent actions	easy	more challenge

Outlines

- Weak Points of Non-Personalized RS
- Preferences and Ratings
- Predictions and Recommendations
- Scoring and Ranking
- Basic Similarity Measures

Asst. Prof. Dr. Rachsuda Setthawong

How to Generate Prediction/Recommendation?



Asst. Prof. Dr. Rachsuda Setthawong

Prediction vs Recommendation

Prediction

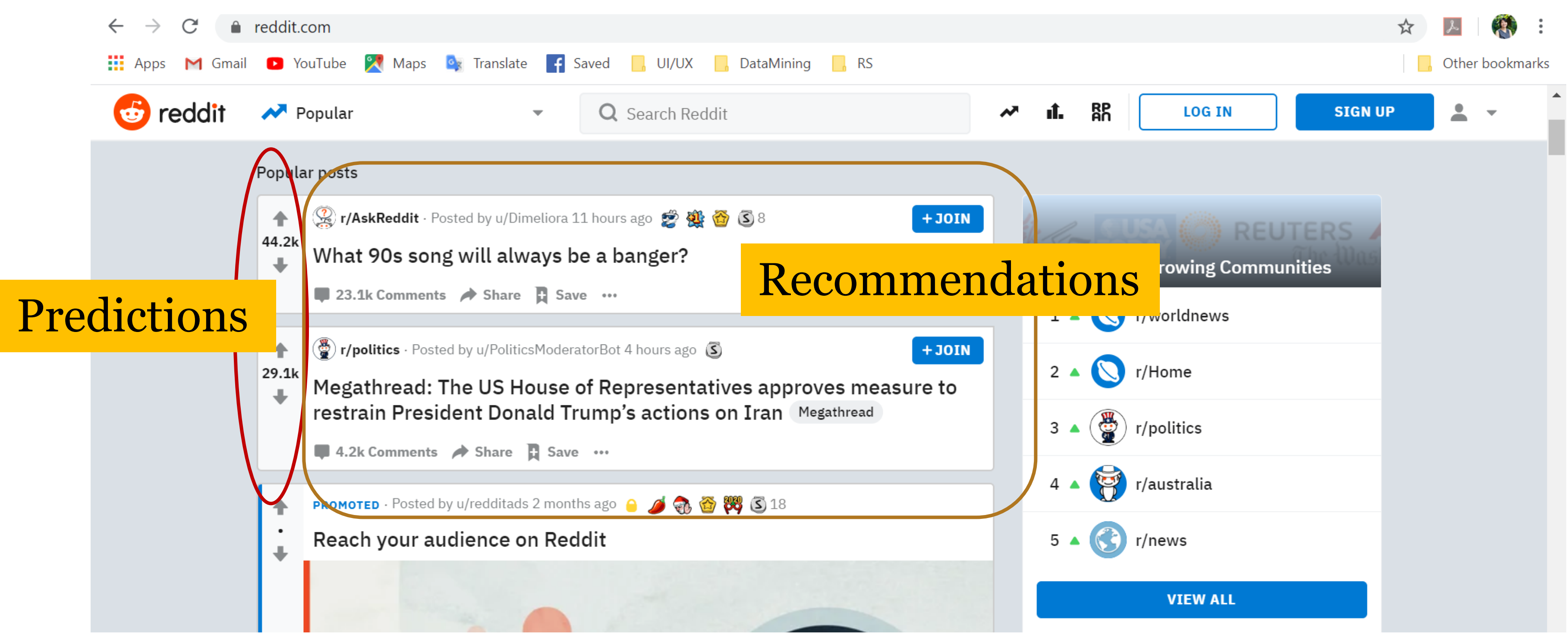
- Estimate how much target users will like an item (a numerical value).

Recommendation

- Rank items based on how much target users will like an item. (e.g., top 10 greatest hits albums)
- Simplify as 'shown' items

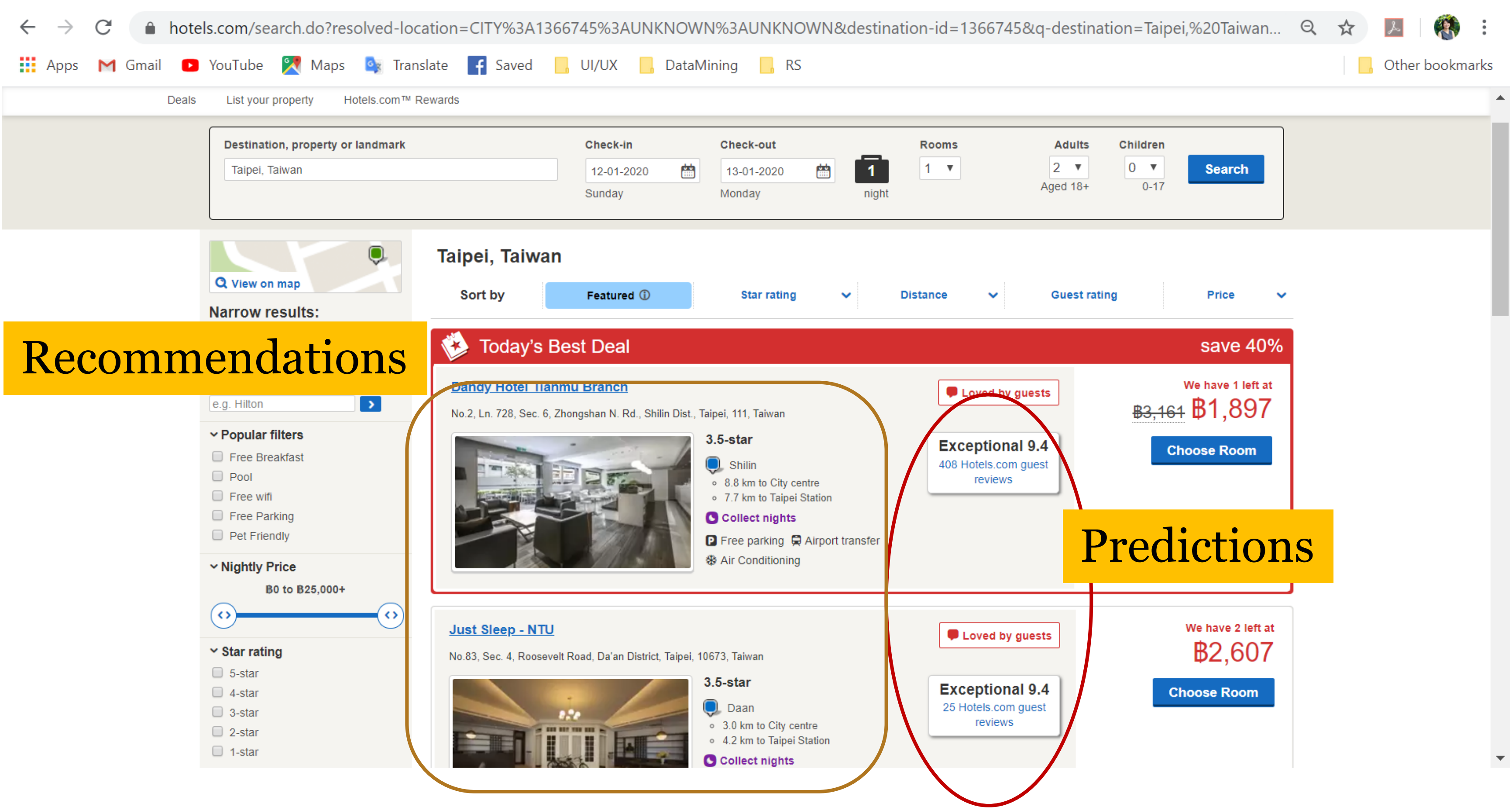
Asst. Prof. Dr. Rachsuda Setthawong

Prediction vs Recommendation



Asst. Prof. Dr. Rachsuda Setthawong

Prediction vs Recommendation

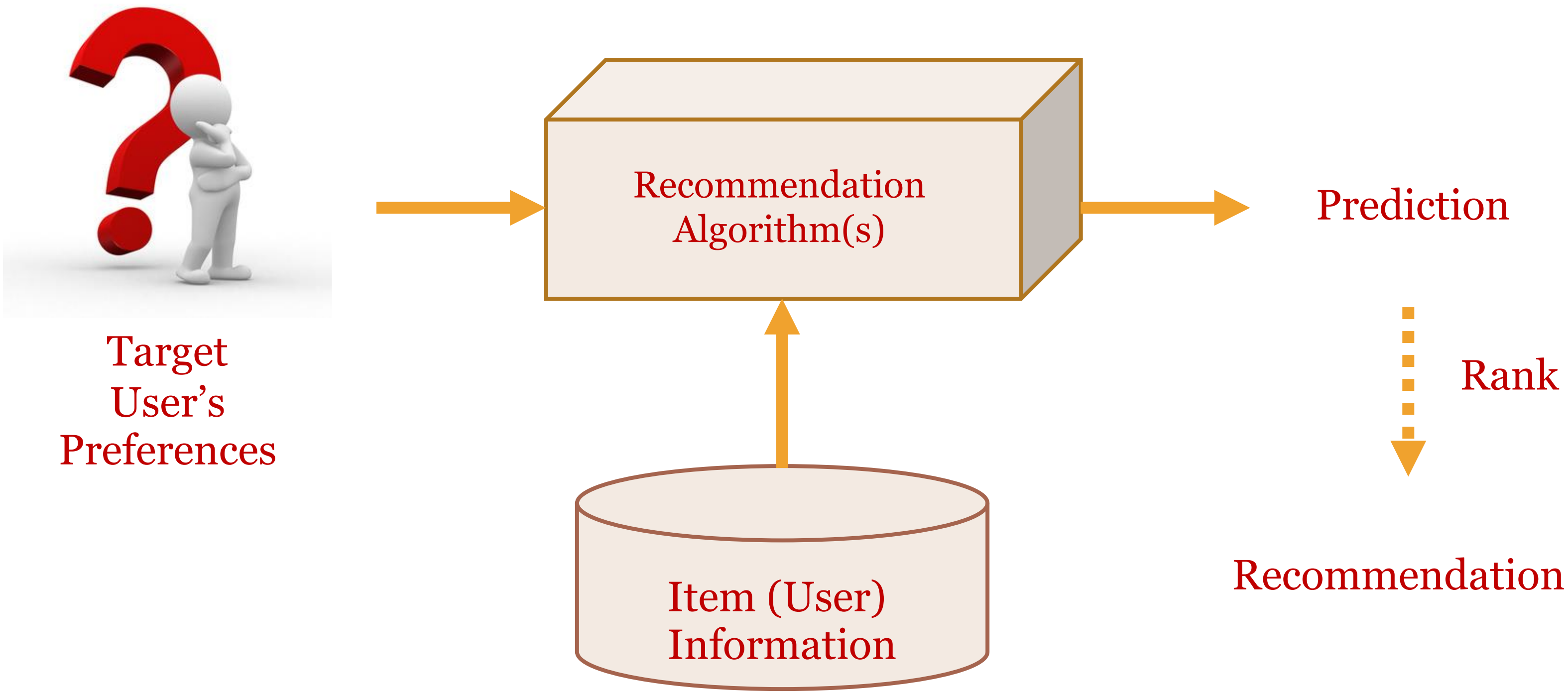


Strong and Weak Points

Predictions	Recommendations
+ helps quantify item	+ provides good choices as a default
- provides something empirical	- Less explore if top-n are not attractive

Asst. Prof. Dr. Rachsuda Setthawong

How to Generate Prediction/Recommendation?



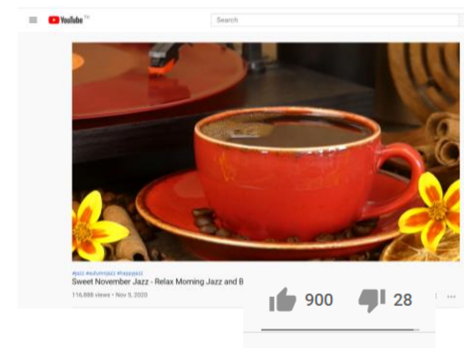
Outlines

- Weak Points of Non-Personalized RS
- Preferences and Ratings
- Predictions and Recommendations
- Scoring and Ranking
- Basic Similarity Measures

Asst. Prof. Dr. Rachsuda Setthawong

Displaying Aggregating Preferences

- Simple scoring
  - Number of upvotes (likes)
  - Average rating or upvote proportion
  - Percentage of >= 4 stars ('positive')
- Full distribution



Note:
Aggregating preference (Predict)
Rank items (Recommend)

Ranking Approaches

- Rank by predictions
- Rank by frequency/quantity
- Rank by timing
- Rank by domain or business objectives
- Etc.

Ranking Considerations

- Confident levels (confidence on goodness of an item)
- High-risk, high-reward or conservative recommendation
- Domain and business considerations in terms of
    - Lifetime period

    - Business objectives

Pros and Cons of Mean

Pros	Cons
• Present an overall picture of community's opinion • Widely used • Calculate easily	• Few ratings affects low confidence. • Not reflect opinion of a niche controversial group

Asst. Prof. Dr. Rachsuda Setthawong

Cons and Solutions of Mean

- Cons: few ratings affects low confidence.
  - Solutions:
    - Scoring of every item is originally average.
    - Ratings will be adjusted to non-averageness wrt to user preferences.
    - *k (a damping term)* controls **strength** of evidence required.
    - μ is **global mean**.

$$damped\_mean\ (i) = \frac{\sum_u r_{ui} + k\mu}{n + k}$$

- Cons: not reflect opinion of a niche controversial group
  - Solutions: Personalization

Additional material: https://medium.com/@tomar.ankur287/non-personalised-recommender-system-in-python-42921cd6f971

Asst. Prof. Dr. Rachsuda Setthawong

An Example

item 'i' **global** mean

$$damped\_mean\ (i) = \frac{\sum_u r_{ui} + k\mu}{n + k}$$

$|u_i|$

a damping term

- Suppose that k = 5
  - μ = 19/7 = 2.714
  - damped_mean ('**Zootopia**')

    =[(2+2+3) + (5×2.714)] / (3+5)

    = 2.57

User	Movie	Rating
Ann	**Zootopia**	**2**
Pete	**Zootopia**	**2**
Kate	**Zootopia**	**3**
Ann	Mona	2
Pete	Mona	3
Ann	Big Hero	4
Kate	Big Hero	3

Notice: when there are a few ratings, it will damp some extreme positive ratings.

Note. If using a simple mean, the calculated rating is 2.33.

Δ = 2.57 - 2.33 = +0.2

Asst. Prof. Dr. Rachsuda Setthawong

An Example (Cont.)

User	Movie	Rating
u1	**Zootopia**	**4**
u2	**Zootopia**	**3**
u3	**Zootopia**	**2**
u4	**Zootopia**	**3**
u5	**Zootopia**	**3**
u6	**Zootopia**	**2**
u7	**Zootopia**	**4**
u8	**Zootopia**	**2**
u9	**Zootopia**	**3**
u10	**Zootopia**	**3**
Ann	**Zootopia**	**2**
Pete	**Zootopia**	**2**
Kate	**Zootopia**	**3**
Ann	Mona	2
Pete	Mona	3
Ann	Big Hero	4
Kate	Big Hero	3

- Suppose that k = 5
  - μ = 48/17 = 2.824
  - damped_mean ('Zootopia')
    = (36 + (5*2.824))/ (13+5)
    = 2.784

Notice: when the number of rating increases, the damping factor has less effect.

Note. If using a simple mean, the calculated rating is 2.769.
Δ = 2.784 - 2.769 = +0.015

Asst. Prof. Dr. Rachsuda Setthawong

Hacker News Ranking Algorithm

Score = (P-1) / (T+2) ^{G}

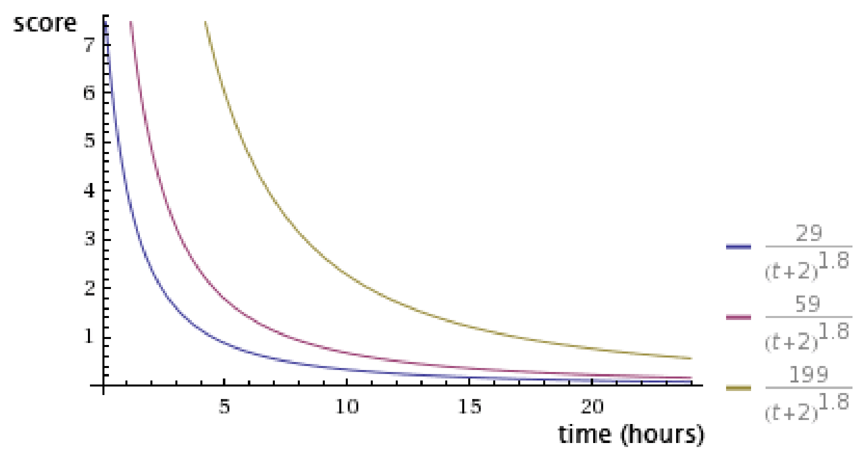
How score is behaving over time

where,

P = points of an item (upvote – downvote)

T = time since submission (in hours)

G = Gravity (defaults is 1.8.)



Q: According to the graph given, when will the score of an item decrease to 1?

Image Source: https://medium.com/hacking-and-gonzo/how-hacker-news-ranking-algorithm-works-1d9b0cf2c08d#.zfcbvz4f8

The Default Story Algorithm in Reddit (Hot Ranking)

$$f(t_s, y, z) = \underbrace{\log_{10} z}_{\text{Factor of \#upvote}} + \underbrace{(y t_s / 45000)}_{\text{Factor of time (aging)}}$$

t_s = Time (in seconds) since Reddit epoch,

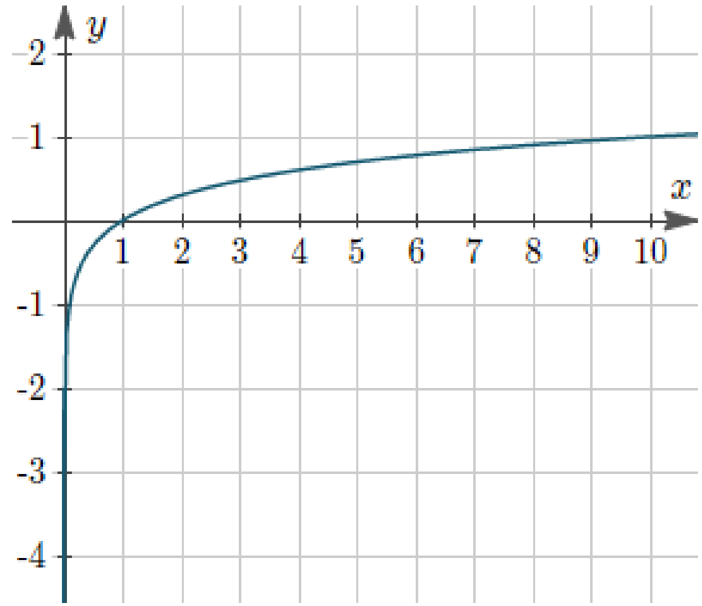
epoch = datetime(1970, 1, 1)

x = #upvotes – #downvotes

$$z = \begin{cases} x \ \ if \ x \geq 1 \\ 1 \ if \ x < 1 \end{cases}$$

$y \in \{-1, 0, 1\}$,   y denotes a signed function

$$y = \begin{cases} 1 \ if \ x > 0 \\ 0 \ if \ x = 0 \\ -1 \ if \ x < 0 \end{cases}$$



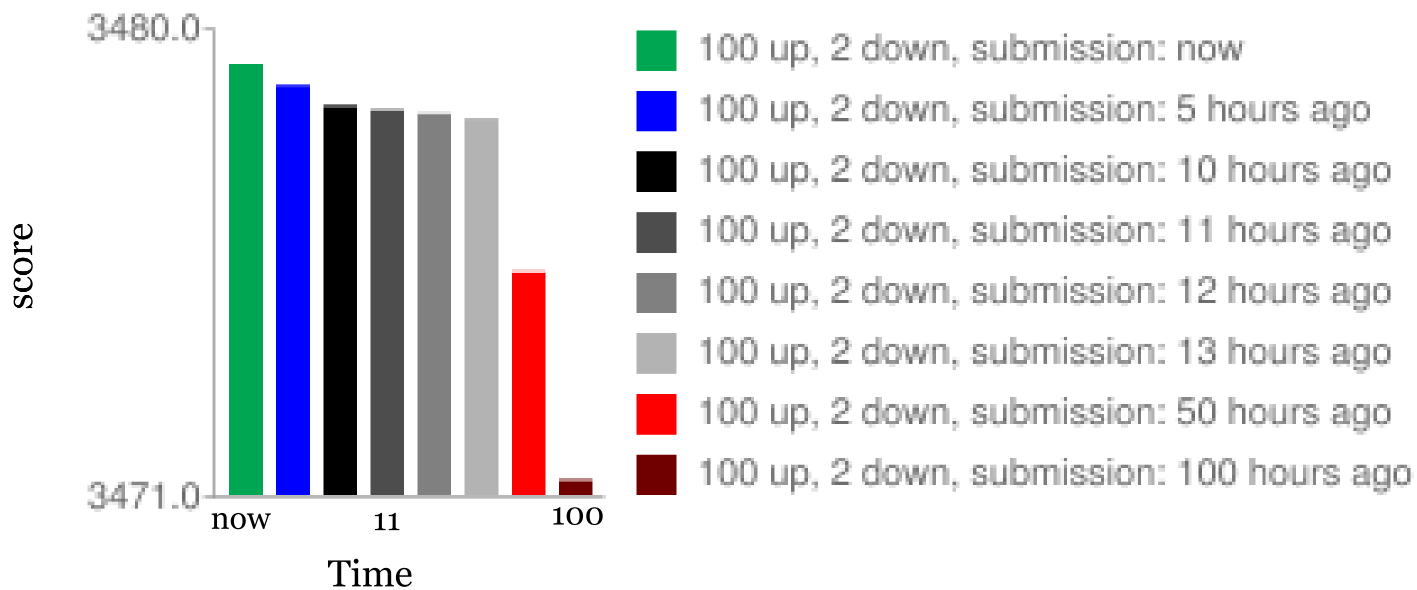
The graph of $y = \log_{10}(x)$

Effect of #upvote

Source: https://medium.com/hacking-and-gonzo/how-reddit-ranking-algorithms-work-ef111e33d0d9#.bj4fokhfm

Asst. Prof. Dr. Rachsuda Setthawong

Effects of Submission Time



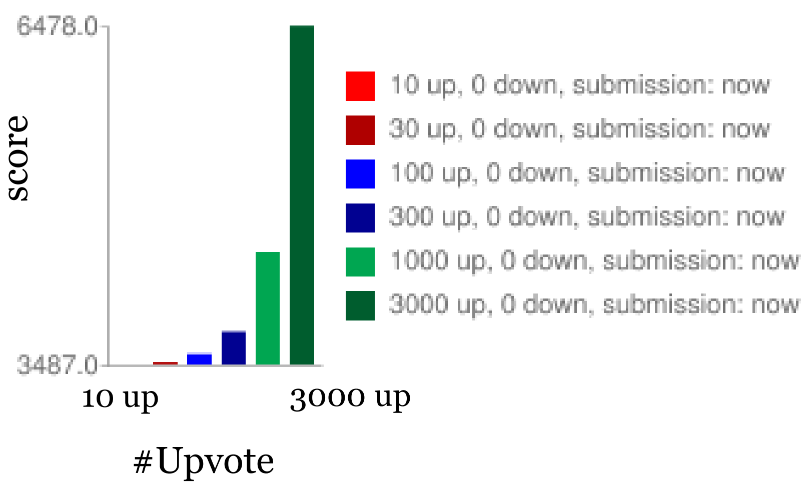
- Submission time is significant to ranking (the newer stories the higher rank is.)
- Newer stories has a higher score than older.

Source: https://medium.com/hacking-and-gonzo/how-reddit-ranking-algorithms-work-ef111e33d0d9#.bj4fokhfm

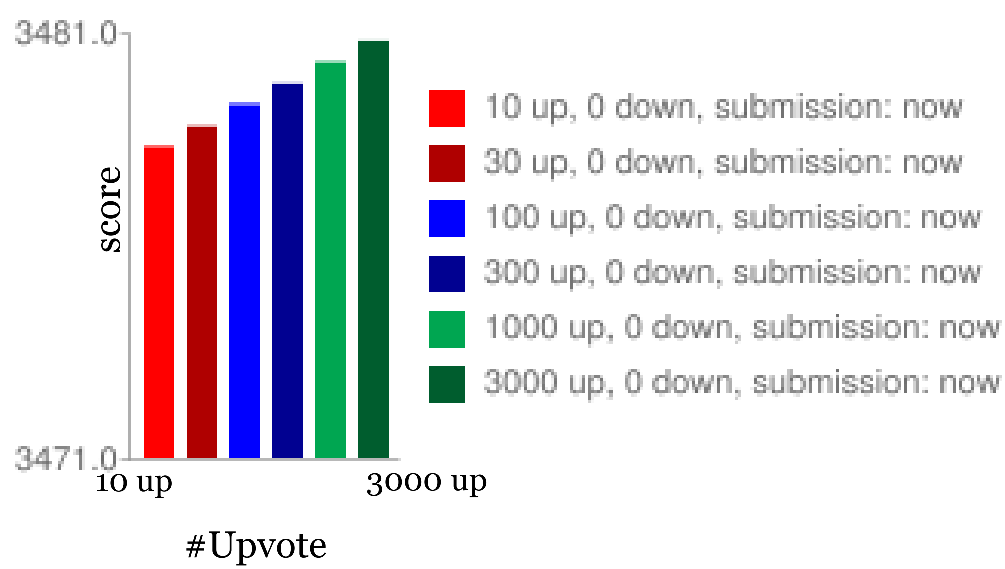
Asst. Prof. Dr. Rachsuda Setthawong

The Effect of Using Logarithm Scale on The Calculated Point (Votes)

Without using the logarithm scale



Using the logarithm scale



Source: https://medium.com/hacking-and-gonzo/how-reddit-ranking-algorithms-work-ef111e33d0d9#.bj4fokhfm

Asst. Prof. Dr. Rachsuda Setthawong

More about Reddit

- How Reddit ranking algorithms work
  - https://medium.com/hacking-and-gonzo/how-reddit-ranking-algorithms-work-ef111e33d0d9#.bj4fokhfm
- Reddit dataset:
  - https://www.reddit.com/r/datasets/comments/3mg812/full_reddit_submission_corpus_now_available_2006/

Asst. Prof. Dr. Rachsuda Setthawong

Outlines

- Weak Points of Non-Personalized RS
- Preferences and Ratings
- Predictions and Recommendations
- Scoring and Ranking
- Basic Similarity Measures

Asst. Prof. Dr. Rachsuda Setthawong

What is Similarity?

- Given 2 feature (attribute) vectors, similarity is a measure how two vectors are similar.

Name\Movie	Zootopia	Superman	Star Trek Beyond	The Angry Bird Movie	Ghost Busters
Ann	Yes		Yes		Yes
Pete	Yes	Yes	Yes	Yes	Yes
Kate		Yes	Yes		
Jason		Yes	Yes	Yes	

- sim(x, y) = ?

Asst. Prof. Dr. Rachsuda Setthawong

Basic Similarity Measures

- Confidence (Association Rule Mining)
  - Asymmetric measure
  - Measure how likely a user is to rate one given that they rated the other
  - **sim(x, y) = (items$_x$ ∩ items$_y$) / items$_x$**
  - Example:
    - sim(Ann, Pete) = 3/3 = 1
    - sim(Pete, Ann) = 3/5 = 0.6
    - sim(Kate, Jason) = 2/2 = 1
    - sim(Jason, Kate) = 2/3 = 0.67

Name\ Movie	Zoo-topia	Super man	Star Trek Be-yond	The Angry Bird Movie	Ghost Bust-ers
Ann	Yes		Yes		Yes
Pete	Yes	Yes	Yes	Yes	Yes
Kate		Yes	Yes		
Jason		Yes	Yes	Yes	

Asst. Prof. Dr. Rachsuda Setthawong

Jaccard Coefficient

- Measure the overlap that *x* and *y* share with their attributes.
- **$J = M_{11} / (M_{01} + M_{10} + M_{11})$**
- where,
  - M_{11} = the total number of attributes where *x* and *y* both have a value of 1.
  - M_{01} = the total number of attributes where the attribute of *x* is 0 and the attribute of *y* is 1.
  - M_{10} = the total number of attributes where the attribute of *x* is 1 and the attribute of *y* is 0.

Asst. Prof. Dr. Rachsuda Setthawong

Jaccard Coefficient's Example

- sim(Ann, Pete) = sim(Pete, Ann) = 3/5 = 0.6
- sim(Kate, Jason) = sim(Jason, Kate) = 2/3 = 0.67
- *Equivalent to*
  - $sim(x, y) = (items_x \cap items_y) / (items_x \cup items_y)$

Name\ Movie	Zoo-topia	Super man	Star Trek Be-yond	The Angry Bird Movie	Ghost Bust-ers
Ann	Yes		Yes		Yes
Pete	Yes	Yes	Yes	Yes	Yes
Kate		Yes	Yes		
Jason		Yes	Yes	Yes	

Practice 2-1

1. Create a dataset D1 of all students in class with ratings of the 3 selected movies (scaling 1-5)
2. Calculate scores of the 3 movies in D1 using Damped Mean
3. Rank items based on the ranking method specified in Q2.

4. Link to edit the dataset D1 is provided in MS Teams' channel.

Asst. Prof. Dr. Rachsuda Setthawong

Practice 2-2

1. Create a dataset D2 of all students in class with upvote (U) or downvote (D)

2. Calculate scores of the 5 topics in D2 using

   Hacker News Ranking Algorithm
3. Rank items based on the ranking method in Q1.

1. Link to edit the dataset D2 is provided in MS Teams' channel.