DATA ANALYTICS

Unit 4: Introduction to Recommendation System

Jyothi R.
Department of Computer Science and Engineering

DATA ANALYTICS

Why Recommendation Systems?

"We are leaving the age of information and entering the age of recommendation"

-Chris Anderson in "The Long Tail"

DATA ANALYTICS
Age of Recommendation

Search:

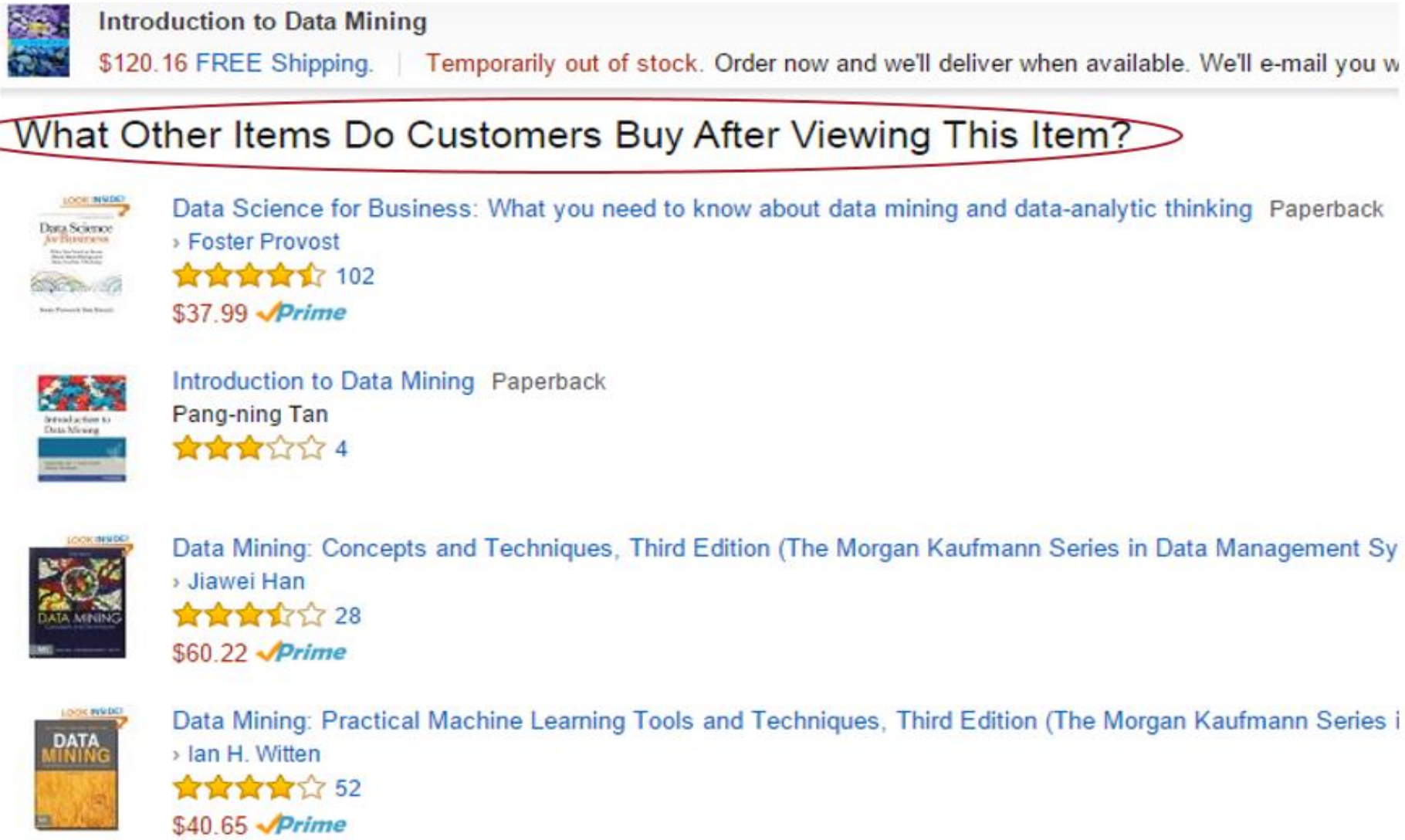
User ➡ Items

Recommend:

Items ➡ User

DATA ANALYTICS
Amazon A personalized online store

DATA ANALYTICS
Amazon A personalized online store



DATA ANALYTICS
Recommender Problem

A Good recommender

- Show programming titles to a software engineer and baby toys to a new mother
- Don’t recommend items, which user already knows or would find anyway.
- Expand User’s taste without offending or annoying him/her..

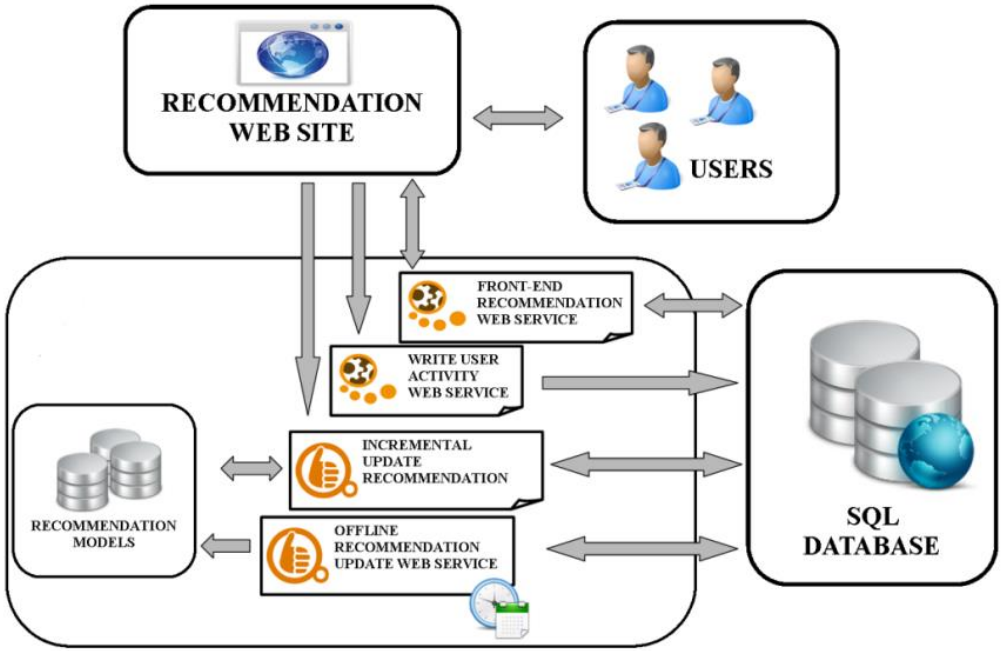
Challenges

- Huge amounts of data, tens of millions of customers and millions of distinct catalog items.
- Results are required to be returned in real time.
- New customers have limited information.
- Customer data is volatile.

DATA ANALYTICS
Amazon's solution

1. Amazon Recommendation Engine: Amazon's model that implements recommendation algorithm. Recommendation algorithm is designed to personalize the online store for each customer.

2. Recommendation Engine Workflow:



DATA ANALYTICS

Recommendations:

Introduction to Recommendation system

- Recommendation system - problem of information filtering
- Enhance user experience
  - Assist users in finding information
  - Reduce search and navigation time
- Recommender systems are the most popular applications of data science today, to Increase productivity, Increase credibility.
- They are used to predict the “rating” or “preference” that a user would give to an item.
- Amazon uses it to suggest products to customers.
- YouTube uses it to decide which video to play next on auto play, and,
- Facebook uses it to recommend pages to like and people to follow.
- Most of the companies business model and its success revolves around the potency of their recommendations.

Goals of Recommender Systems

1. **Prediction version of problem:** the first approach is to predict the rating value for a user-item combination. It is assumed that training data is available, indicating user preferences for items. For m users and n items, this corresponds to an incomplete mxn matrix, where the specified values are used for training.

2. **Ranking version of problem:** In practice, it is not necessary to predict the ratings of users for specific items in order to make recommendations to users. The determination of the top-k items is more common than the determination of top-k users.

Goals of Recommender Systems Contd.

In order to achieve broader business-centric goal of increasing revenue,

the operational and technical goals of recommender systems are as follows

1. **Relevance**: Users are more likely to consume items they find interesting, rating value for a user-item combination.
2. **Novelty:** Recommender systems are truly helpful when the recommended item is something that the user has not seen in the past. For example, Popular movies of a preferred genre would rarely be novel to the user.
3. **Serendipity:** The items recommended are somewhat unexpected, and therefore there is a modest element of lucky discovery. Recommendations are truly surprising to the user. It leads to sales diversity or beginning a new trend of interest in the user.
4. **Increasing recommendation diversity:** It has the benefit of ensuring that the user does not get bored by repeated recommendation of similar items.

Goals of Recommender Systems Contd.

- Aside from these concrete goals, a number of soft goals are also met by the recommendation process both from the perspective of the user and merchant.
- The broad diversity of recommender systems that were built either as research prototype, or are available today as commercial systems in various problem settings

1. GroupLens Recommender System
2. Amazon.com Recommender System
3. Netflix Movie Recommender System
4. Google News Personalization System
5. Facebook Friend Recommendations

DATA ANALYTICS

The Spectrum of Recommendation Applications

1. Collaborative Filtering Models.
    i) Memory - based collaborative filtering
        a) User-Based collaborative filtering
        b) Item-based collaborative filtering
    ii) Model-Based Methods
- Types of Ratings. Implicit and Explicit Ratings
- Relationship with missing values.

2. Content-Based Recommender systems
3. Knowledge-Based Recommender Systems
    i)Constraint-based recommender systems:
    ii)Case-based recommender systems:
4. Demographic Recommender systems
5. Hybrid and Ensemble-Based Recommender Systems

Gathering Ratings

Types of Ratings

- Explicit
- Ask people to rate items
- Doesn’t scale: only a small fraction of users leave ratings and reviews

- Implicit
- Learn ratings from user actions
- E.g., purchase implies high rating
- What about low ratings?

Domain_specific Challenges in Recommender Systems

1. Context-Based Recommender Systems: It could include time, location, or social data. For example, the types of clothes recommended by a retailer might depend both on the season and location of the customer. Even particular type of festival or holiday affects the underlying customer activity.
2. Time-Sensitive Recommender Systems

i. The rating of an item might evolve with time, as community attitudes evolve and the interests of users change over time. User interests, likes, dislikes, and fashions inevitably evolve with time.

ii. The rating of an item might be dependent on the specific time of day, day of week, month, or season.

iii. For example, it makes little sense to recommend winter clothing during the summer, or Raincoats during the dry season.

Domain_specific Challenges in Recommender Systems

3. Location-Based Recommender Systems
    i) User-Specific Locality
    ii)Item-specific Locality
4. Social Recommender Systems
    i).Structural Recommendation of Nodes and Links
    ii)Product and Content Recommendations with social influence.
    iii)Trustworthy Recommender Systems
    iv)Leveraging Social Tagging Feedback for Recommendations

DATA ANALYTICS

Advanced Topics and Applications

1. The Cold-Start Problem in Recommender Systems: Most people have not rated most items

Cold Start:

- New items have no ratings
- New users have no history

2. Attack-Resistant Recommender Systems
3. Group Recommender Systems
4. Multi-Criteria Recommender Systems
5. Active Learning in Recommender Systems
6. Privacy in Recommender Systems
7. Application Domains

From Scarcity to Abundance

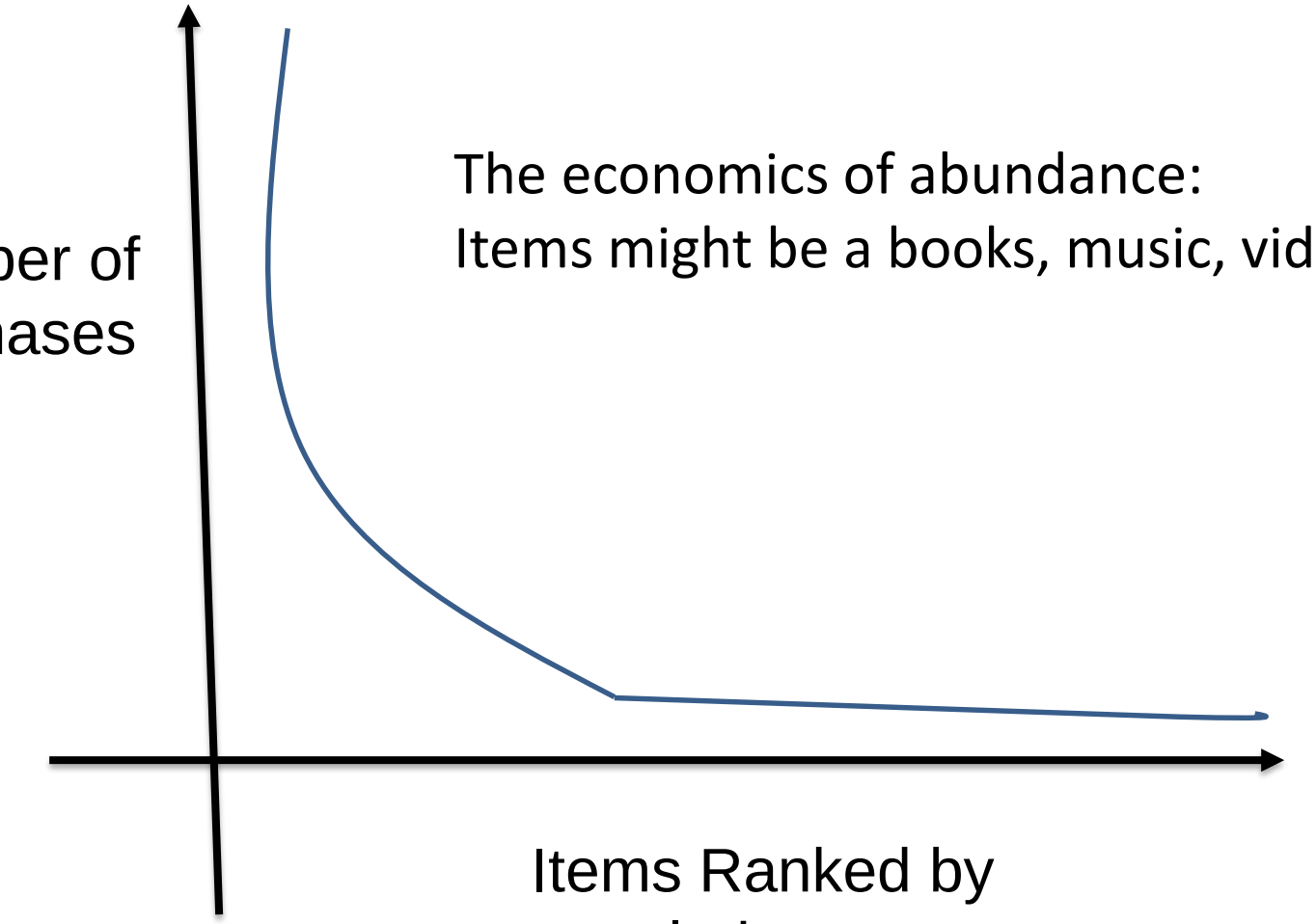
- Shelf space is a scarce commodity for traditional retailers
  TV networks, Movie Theatres.

- More choice necessitates better filters.

- The web enables near-zero-cost dissemination of information about products.
  From scarcity to abundance
  Gives rise to the “Long Tail” phenomenon.

DATA ANALYTICS

The long tail(1):



The economics of abundance:
Items might be a books, music, videos, or news articles

DATA ANALYTICS

Overview of recommendations

1. Editorial and hand curated i) List of favourites ii) List of “essential” items
2. Simple aggregates. i) Top 10, Most Popular, Recent Uploads.
3. Tailored to individual users. i) Amazon, Netflix, Pandora's

Formal Model

Utility Function : A function that looks at every pair of a customer and item and maps it.

U: CXS→ R

I. C= set of customers and
II. S=Set of items
III. R=Set of ratings, it is a totally ordered set eg.-5 stars, real no in[0..1]

Utility Matrix

	Avatar	KGF	Matrix	Bahubali
Alice	1		0.2	
Bob		0.5		0.3
Carol	0.2		1	
David				0.4

Key Problems

1. Gathering “Known” ratings for matrix: How to collect the data in the utility matrix
2. Extrapolate unknown ratings from the known ones: mainly interested in high unknown ratings. i.e. we are not interested in knowing what you don’t like but what you like.
3. Evaluating extrapolation methods: how to measure success/performance of recommendation methods.

References

Text Book:

"Business Analytics, The Science of Decision Making", U. Dinesh Kumar, Wiley 2017

"Recommender Systems, The text book, Charu C. Aggarwal, Springer 2016 Section 1.

Image Courtesy

https://www.scribd.com/presentation/414445910/CS548S15-Showcase-Web-Mining

https://towardsdatascience.com/image-recommendation-engine-leverage-transfert-learning-ec9af32f5239

https://www.youtube.com/watch?v=1JRrCEgiyHM

www.amazon.com

http://elico.rapid-i.com/recommender-extension.html

THANK YOU

Jyothi R.
Assistant Professor,
Department of Computer Science
jyothir@pes.edu