March 17, 2017

Recommender systems

Evolutionary Computation and Machine Learning @ RMIT

INFOREADY

Maksud Ibrahimov

March 2017

Outline

- Who am I?
- What are the recommender systems?
- Characteristics of a good recommender system
- Algorithms, metrics and evaluation

Who am I?

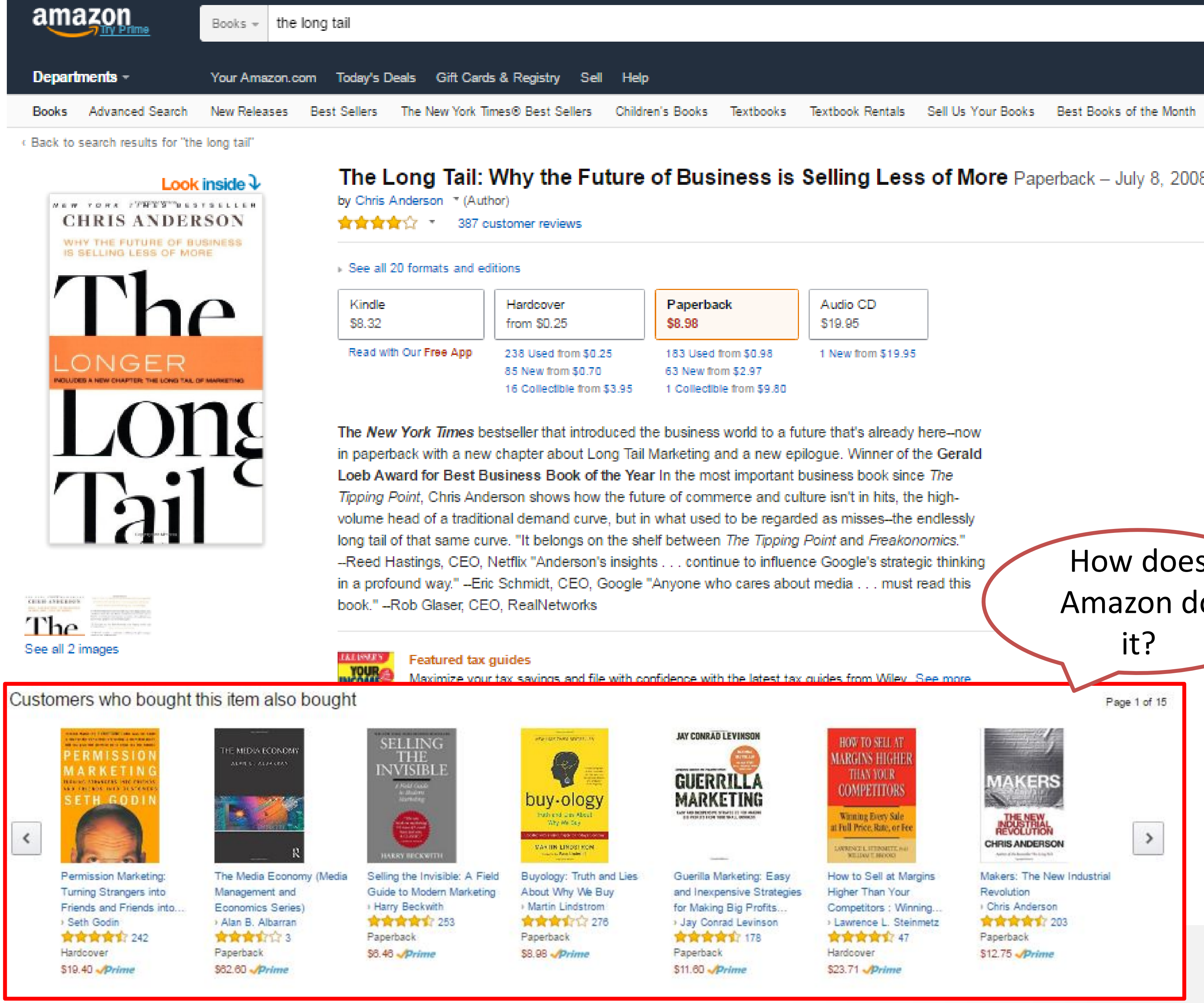
— Chief Data Scientist at InfoReady - a leading Australian data and analytics business
— PhD in Artificial Intelligence from the University of Adelaide
— Ex UN/UNFPA, SolveIT, Schneider Electric, ABB/Ventyx
— Over last 10+ years, have worked and improved operations of major companies in mining, manufacturing, retail, telecom and logistics through machine learning, optimisation and simulation
— Founder of Adaptic Solutions, Australian boutique data science consultancy. Acquired by Infoready in 2015

INFOREADY

INFOREADY

2008


2008
Industries
SolveIT Software
Schneider Electric
ABB
adaptic
INFOREADY
SUB CRUCE LUMEN
THE UNIVERSITY of ADELAIDE
Technologies
hadoop
Spark
INFOREADY



How does Amazon do it?

Significant share of profits generated by recommender systems

Google news: 38% increase in click-through generated from recommended articles

amazon: 35% of sales come from recommendations

NETFLIX: 60% of sales are driven by recommendations

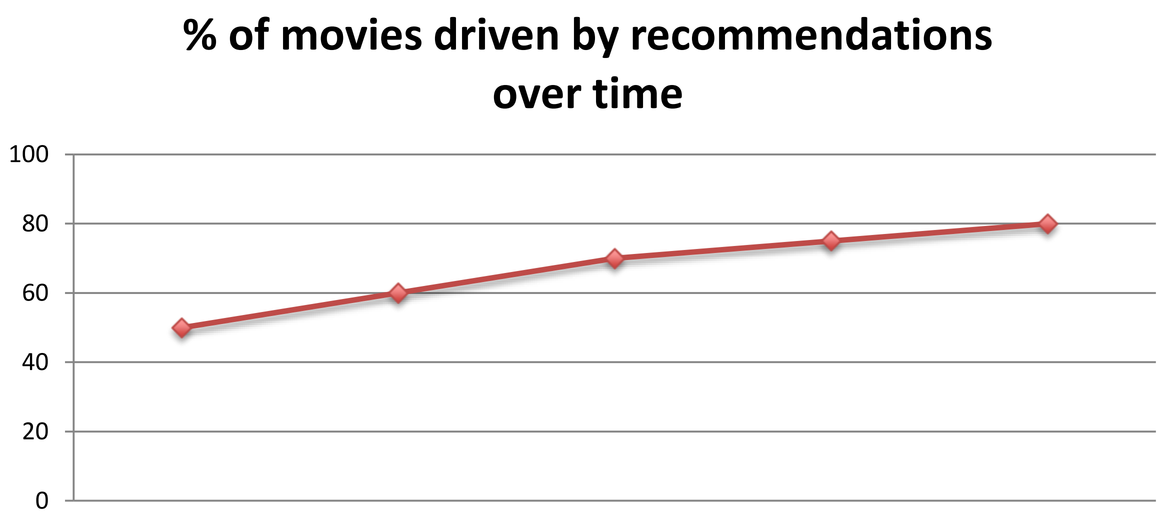
INFOREADY

Source: Amatriain, Xavier (2014), Recommender Systems (Machine Learning Summer School 2014 @ CMU)

In modern age, you don’t find the product, the product finds you

“Search is what you do when you're looking for something. Discovery is when something wonderful that you didn't know existed, or didn't know how to ask for, finds you”

- CNN Money, The race to create 'smart' Google



Source: Netflix

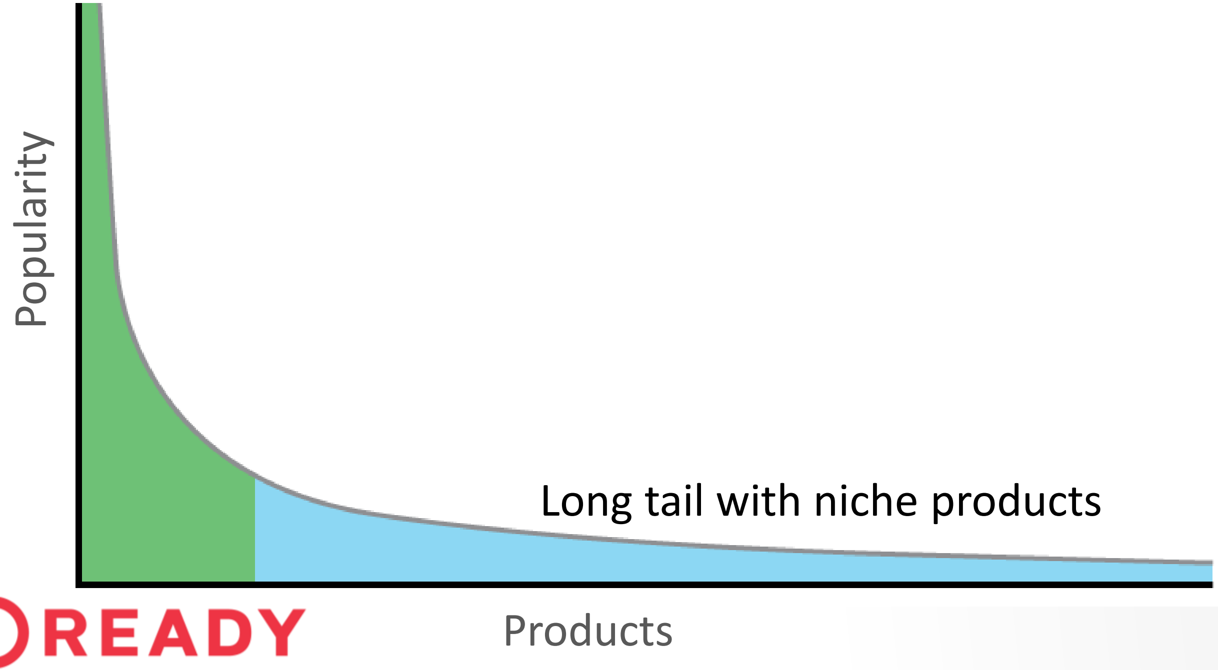
INFOREADY

You don’t find the product, the product finds you

— Search allows you to search by entering the term, **explicitly**

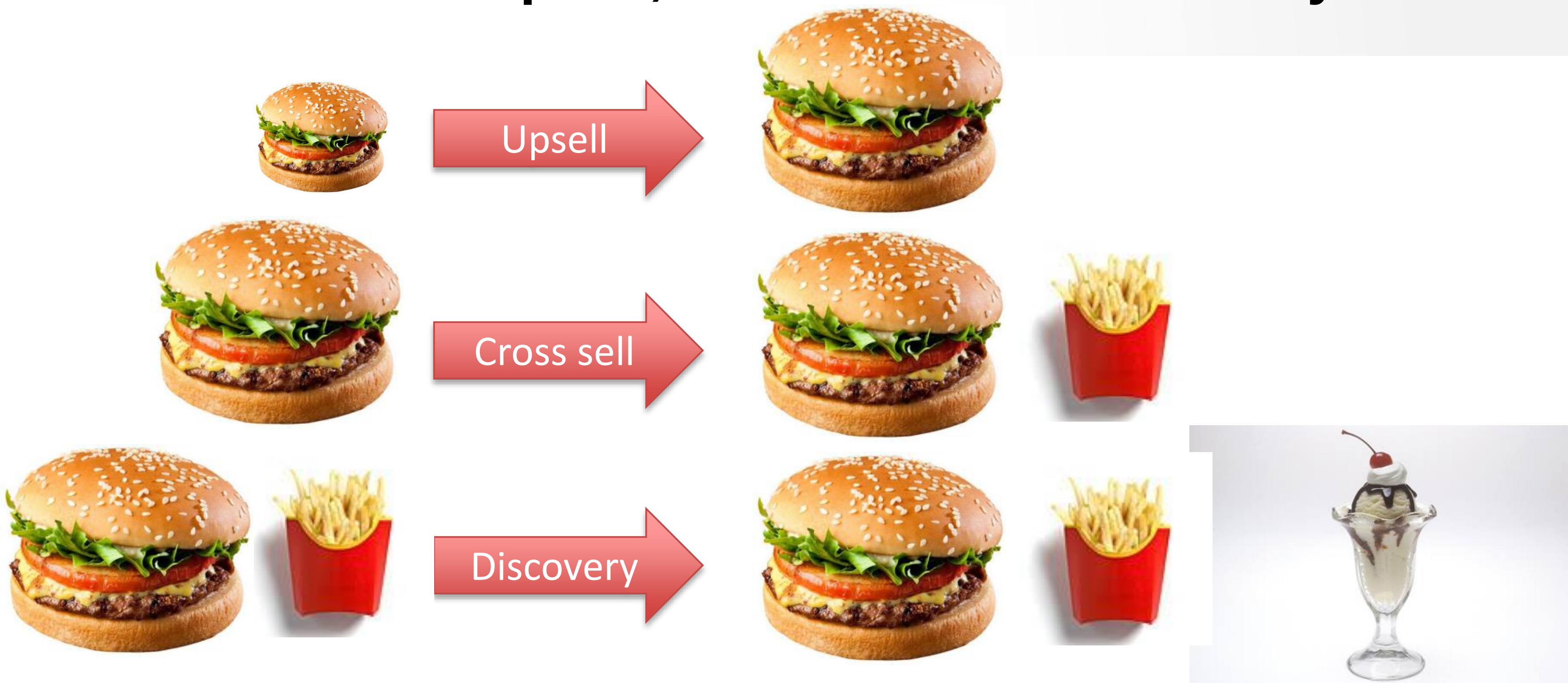
— Recommender systems do not need a search term, it takes it **implicitly**

— Recommender system is a better user experience for finding things than traditional search



INFOREADY

Value add of RS: Upsell, cross sell and discovery



— Retail
— Utilities
— Financial services
— Call centres

Not so much:

- *Transportation*
- *Logistics*
- *Mining*

INFOREADY

Intuition for the recommender problem

Consumers similar to you prefer these offers.

consumers	a population we can see the preferences of
similar to	a comparison which clusters patterns of behaviour
you	an individual, personalised contact
prefer	choose consistently between alternatives given scarcity
these offers	the alternatives (being products, services, content)

Any old recommendation vs. A great recommendation

People like to feel unique, and businesses like offers that people take

— Personalised recommendation tailored based on preferences, qualifications, and previous behaviour

— Each customer is unique and recommendations will be different from other users, certainly within large groups

— We want to find the best recommendation for that person

— Win-win situation

Serendipity – recommendations need to surprise you

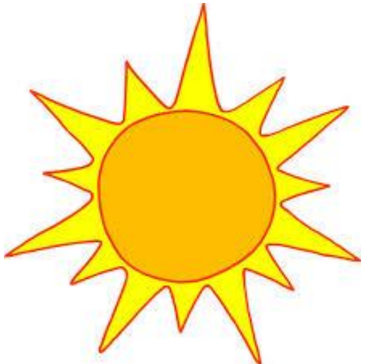
You are definitely going to buy these

Diversity - different product families

I know you
like cheese

INFOREADY

Contextual – differ depending on context

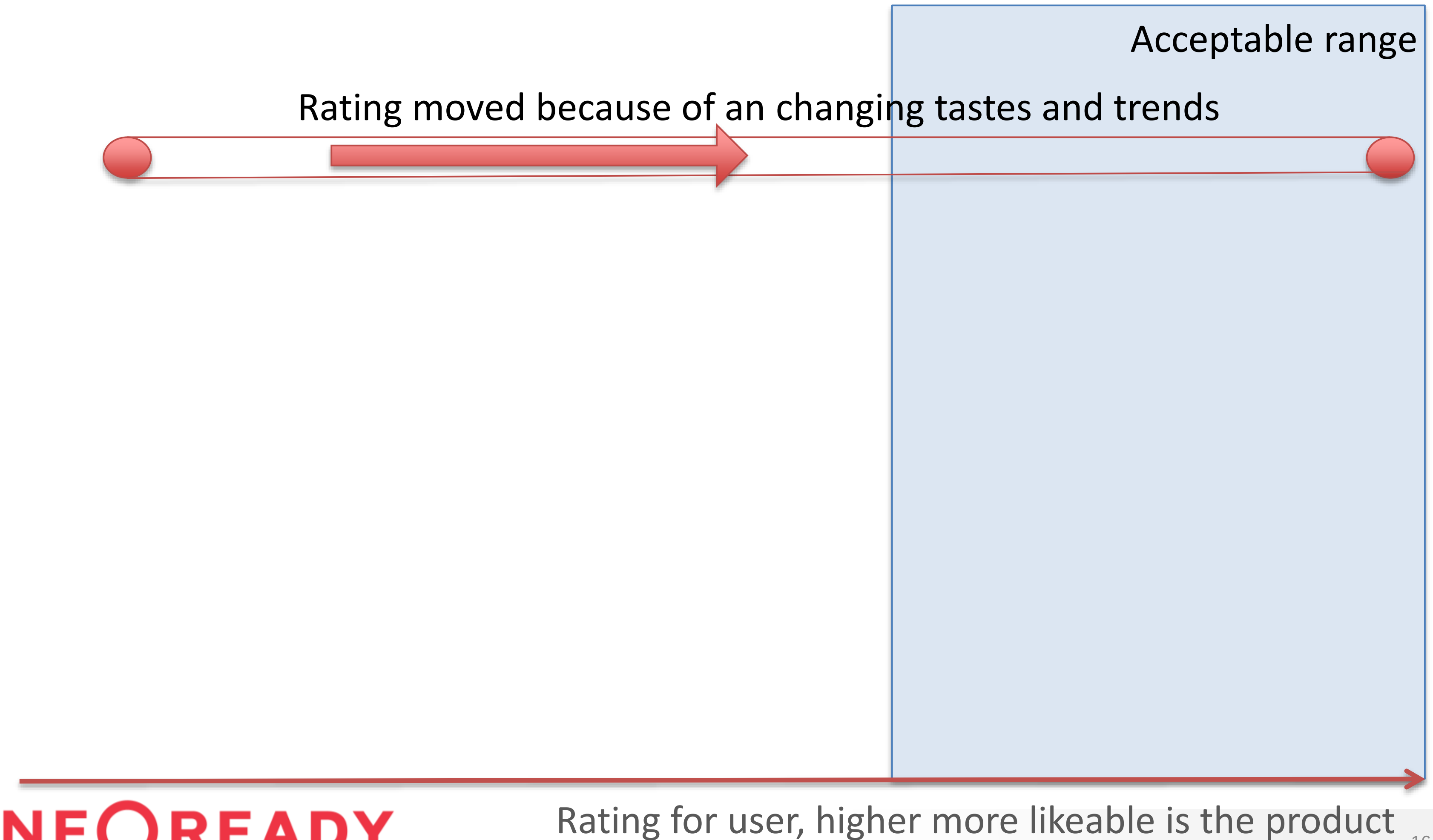


Context: Summer

Context: Winter

INFOREADY

Adaptivity – flexible under changing preferences



INFOREADY

Business value and utility is important

The value the system or user gains from a recommendation

Margin: $1
Rating: 4.96

Margin: $2
Rating: 4.78

INFOREADY

What is a “recommender system”?

— A system that estimates a **utility function** that automatically predicts how a **user** will like an **item** based on the historical data

— **U** is a set of consumers

— **S** is a set of offers

— **F: U x S → R** is a utility function that returns an ordered set

— For each user $u \in U$, choose $s \in S$ that maximises F

— $\forall\ u \in U,\ s_c = \text{argmax}(F(u,s))\ s \in S$

INFOREADY

Define consumers

— "Users" in Netflix terminology
— Identified consumer for targeting offers during interactions
— Attributed preferences
— Additional information: Geo/demographic
  - Called "side information"

INFOREADY

Define offers

— Items in Netflix terminology
  - Movies, books, products, etc.
  - Services
  - Articles

— Uniquely and clearly identify the offer

— Attributed to preferences

— Augmented by additional properties
  - Type of offer and/or hierarchy
  - Hierarchies (eg. merchandise hierarchy) are awesome
  - Colour, size, specs

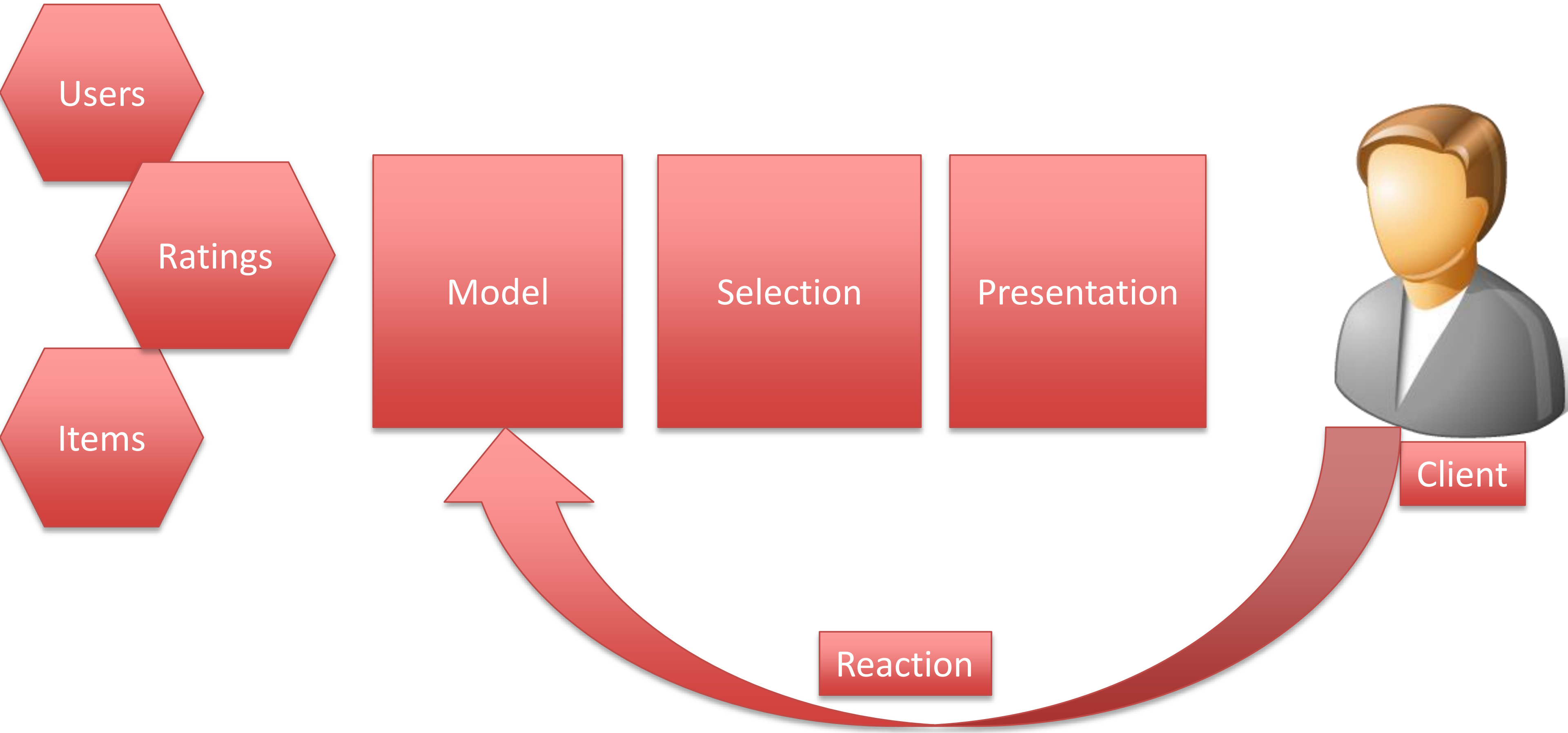
INFOREADY

Define preferences

— Ratings in Netflix terminology
— Preferences is a connection between consumers and offers and allows an ordering/ranking of the offers by a consumer
— Timestamp of the preference helps with adaptivity and context-sensitivity
— Preferences can be stated implicitly or explicitly
  - Explicit preference is a rating, like, review, vote
  - Implicit is click, reading time on page, used in the absence of explicit rating

INFOREADY

Components of a recommender system



INFOREADY

Selected the best recommendations

Given the similar consumers, how do we choose a good recommendation?

— Present N **best** recommendations to the user

— N depends on context – how many can consumer absorb?

— What does best mean?

- Prediction of most preferred is not only consideration

— Needs careful engineering to address recommendation requirements

- Diversity – cluster recommendations and choose a representative from each cluster
- Serendipity – calculate overall popularity of an item and then penalise high popularity items
- Utility – sort by utility function (e.g. margin, sale size)
- Repetition

INFOREADY

Presenting the recommendations, integrating into the business process

- On-line (based on user session)
  - Flash adverts/reminders
  - Promotions/specials
- Mobile app, push or alert, based on location
- During contact
  - during contact-centre interaction
  - while at check-out counter
- Indirectly
  - Inform sales account representatives of offer’s their customers are most likely to be receptive to
- Autonomously
  - Substitutions during stock-out

Types of algorithms for recommender systems. Classic methods

— **Classic methods**

- Popularity based. Good for cold start
- Collaborative filtering – based on similarity (users or items)
- Content based recommenders – based on item features
- Matrix factorisation
- Clustering

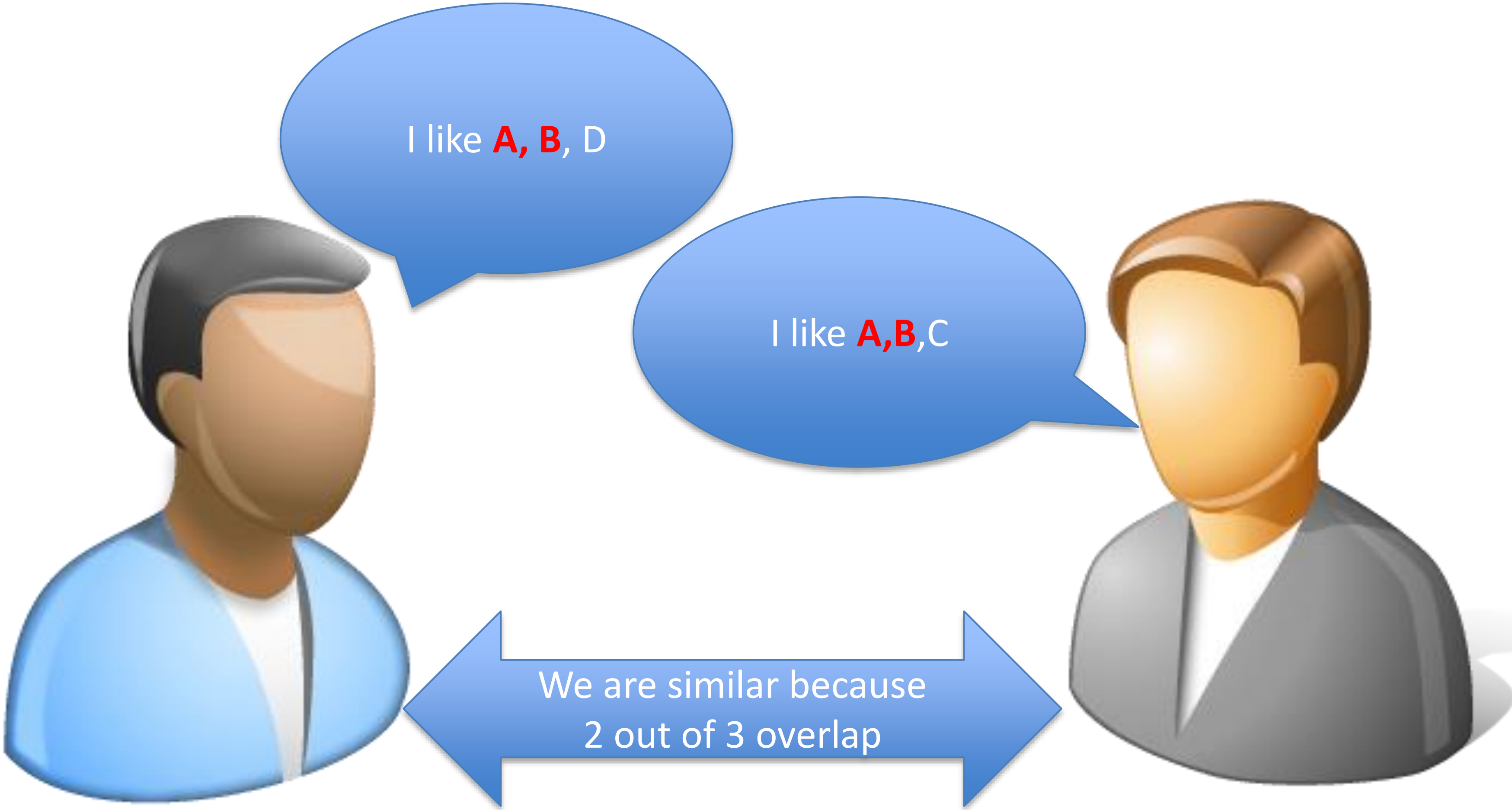
— **Ensembles!**

INFOREADY

Popularity based methods

- Just the most popular item can be a good recommendation as a baseline
- Used for cold start problem, when there is not enough data available about the user or item

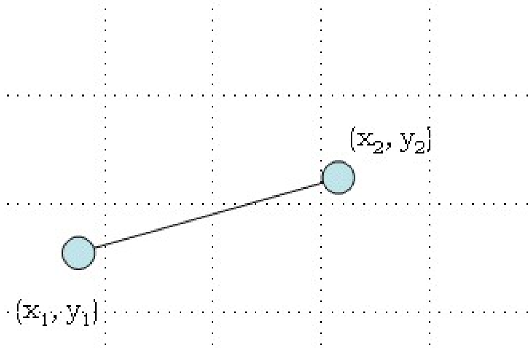
Collaborative Filtering: Intuition for consumer similarity



(This is called Jaccard similarity)

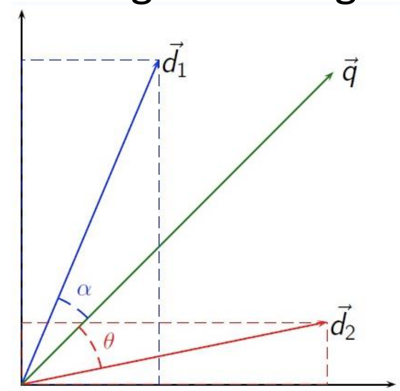
Collaborative Filtering: Similarity functions focus on different things

Euclidean – continuous



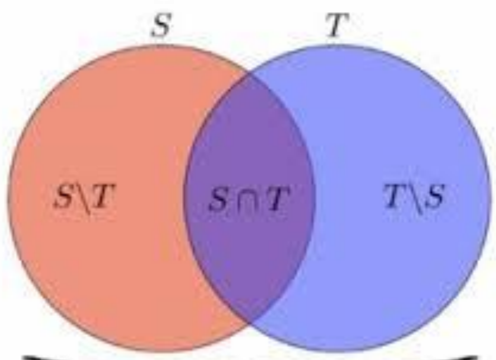
$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

Cosine – ignores magnitude



$$similarity = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

Jaccard – penalises dissimilarity



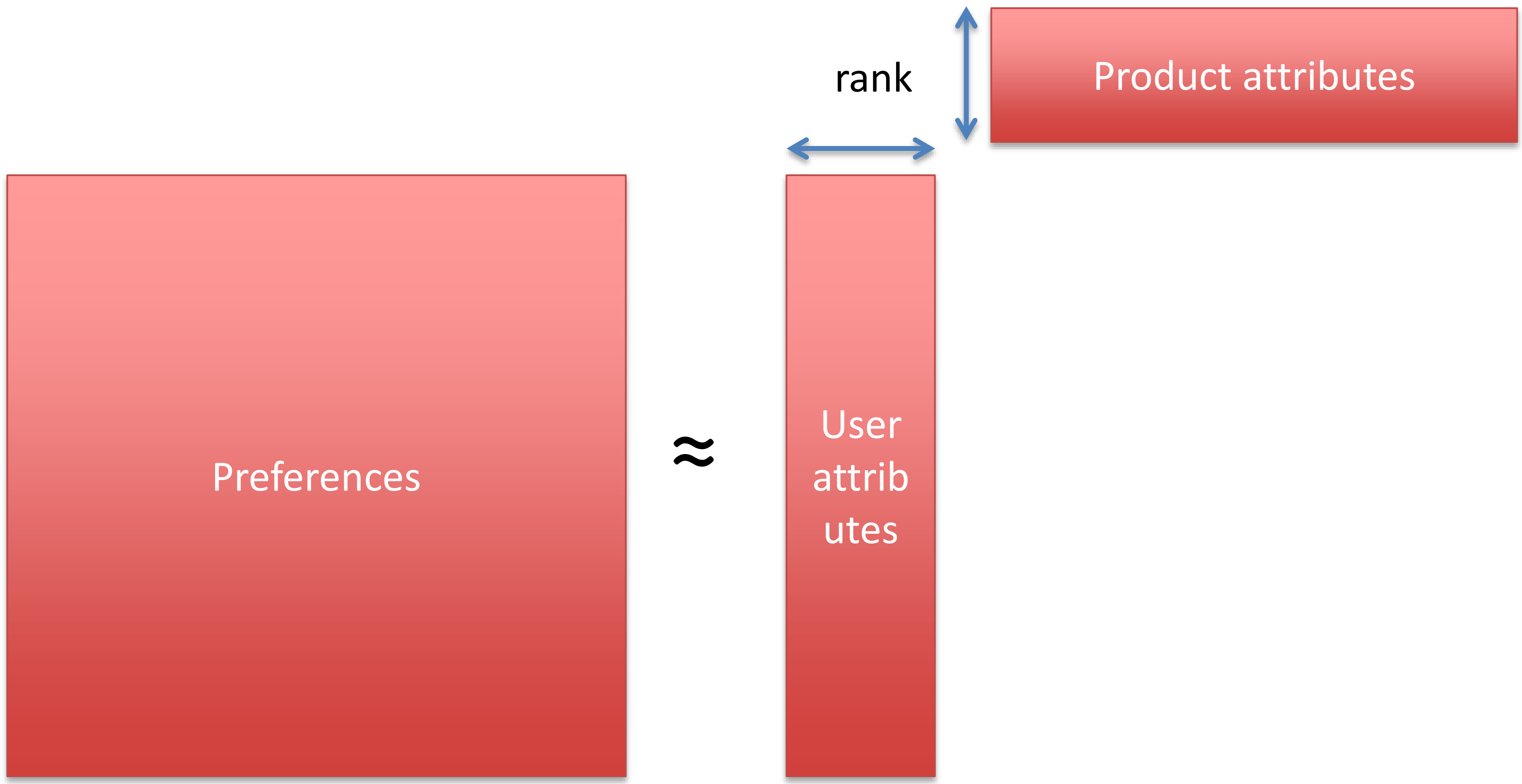
$$J(S,T) = \frac{|S \cap T|}{|S \cup T|}$$

Hamming distance - positional

Gene 1	A	A	T	C	C	A	G	T
Gene 2	T	C	T	C	A	A	G	C
Hamming Distance	1	1	0	0	1	0	0	1

$$D_H = \sum_{i=1}^{k} |x_i - y_i|$$

Matrix factorisation: intuition



Metrics

— Root mean squared error
— Mean absolute error
— Based on top-n list
  - Precision, Recall, Mean Average Precision
  - Number of bad/good items in the list
  - ROC, AUC
— Rank metrics
  - Mean Reciprocal Rank
  - Spearman Rank Correlation
  - Discounted Cumulative Gain
  - Fraction of Concordant Pairs

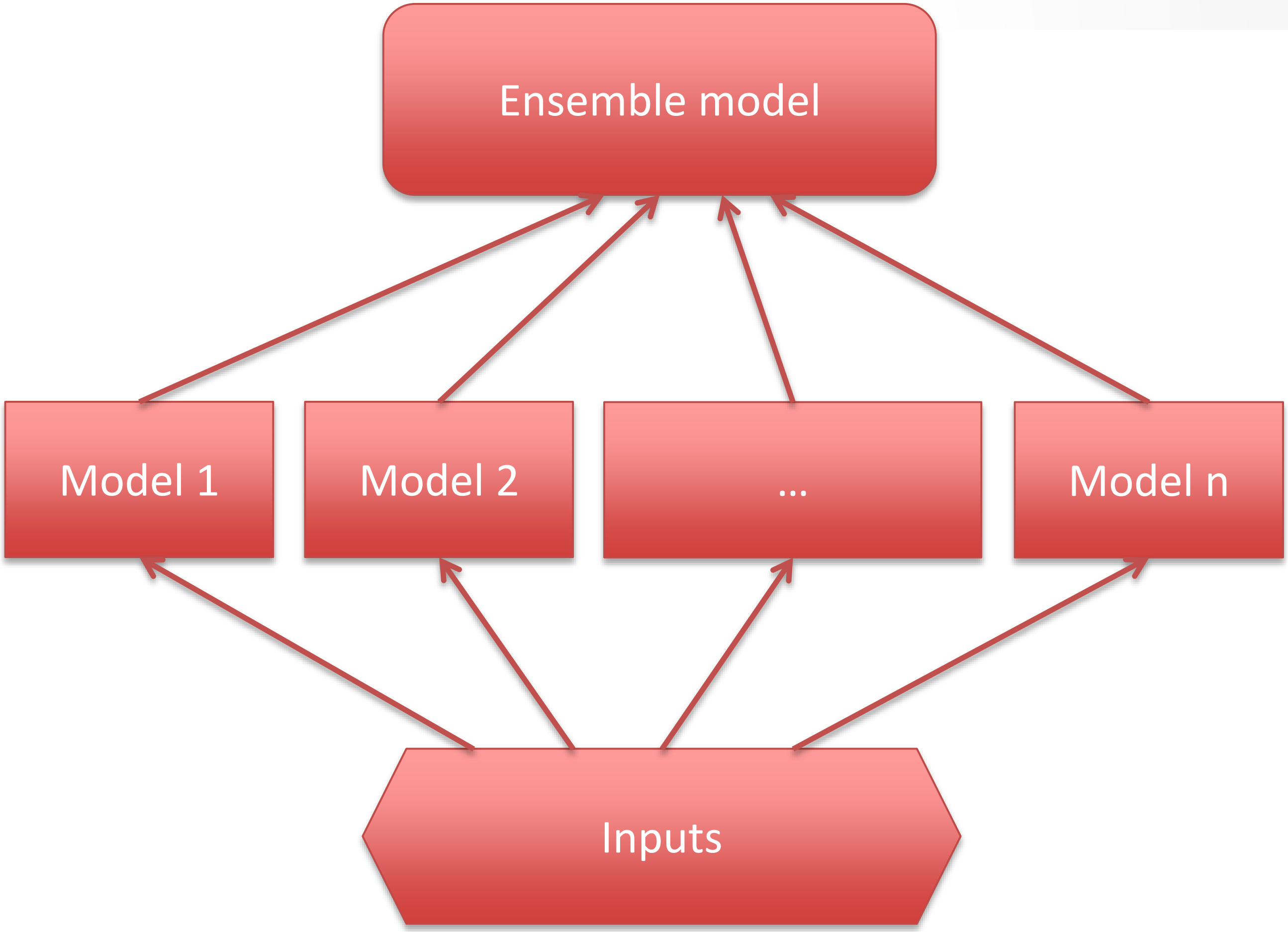
Evaluation

- Cross validation is generally used to evaluate performance of the recommender
- Splits can be done:
  - By ratings
    - Randomly
    - By time
  - By users
    - Randomly
    - By time

Model based

- Clustering
  - Cluster and use popularity based measures within the cluster
- Classifiers
  - SVM, RandomForest, Logistic regression, etc
- Neural networks
  - 1-3 generation NN (perceptrons, backprop, belief networks, etc)
  - Restricted Boltzmann Machines
  - Deep learning

Ensembles improve performance most of the time



INFOREADY

Types of algorithms for recommender systems. Continued.

— Modern methods

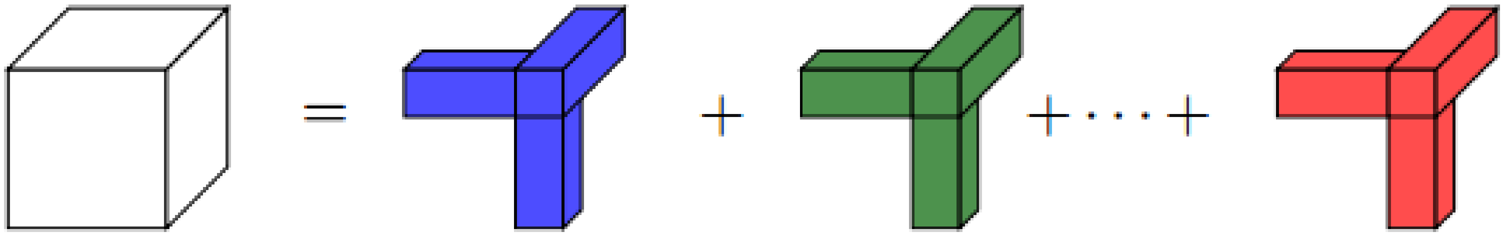
- Association rules (Apriori algorithm)
- Genetic algorithms
- Ranking algorithms
- Tensor factorisation
- Social network based recommendations
- Restricted Boltzmann machines
- Deep learning

INFOREADY

Ranking algorithms

— Pointwise. Regression models
— Pairwise. Minimise number of inversions
— Listwise. Metaheuristics used here

Tensor factorisation



— Tensor-based recommender models push the boundaries of traditional collaborative filtering techniques by taking into account a multifaceted nature of real environments, which allows to produce more accurate, situational (e.g. context-aware, criteria-driven) recommendations.

Thank you!