3.4 Introduction to Recommender Systems

Introduction

- Given a user model (ratings, preferences, ...) and items, find relevance score
- Typically used for ranking

Paradigms of Recommender Systems

- **Personalized recommendations**: user profile & contextual parameters
- **Collaborative**: community data
- **Content-based**: product features
- **Knowledge-based**: knowledge models
- **Hybrid**: combinations of various inputs and/or composition of different mechanism

Collaborative Filtering

Pure CF Approaches

- **Input**: matrix of given user-item ratings
- **Output types**: degree to what the current user will like or dislike a certain item; a top-N list of recommended items

User-based nearest-neighbor CF

Basic algorithm

- given an "active user" u and an item i not yet seen by u
  - find a set of users (peers/nearest neighbors) who liked the same items as u in the past and who have rated item i
  - combine their ratings to predict if u will like item i
- repeat this for all items that u has not seen
- Recommend the best-rated items

Basic assumptions

- if users had similar tastes in the past, they will have similar tastes in the future
- user preferences remain stable and consistent over time

Evaluation

- Business questions lead to empirical evaluation during development and in deployment

Offline Evaluation

- Data: collected in your problem; benchmark datasets
- Remove useless data:
  - **training set**: randomly selected share of known users
  - **testing set**: recommendations based on observed items compared to hidden items

Metrics

- **Precision**: exactness; fraction of relevant items retrieved out of all items retrieved

- **Recall**: completeness; fraction of relevant items retrieved out of all relevant items

  - When tuning system to increase precision, recall decreases

- **F1 Metric**: combine Precision and Recall into a single value

- **Rank Metrics**: extend recall and precision to take the positions of correct items in a ranked list into account

- **Rank Score**: extend recall metric to take the positions of correct items in a ranked list into account

Normalized Discounted Cumulative Gain

- **Discounted cumulative gain (DCG)**: logarithmic reduction factor

- **Idealized discounted cumulative gain (IDCG)**: assumption that items are ordered by decreasing relevance

- **Normalize discounted cumulative gain (nDCG)**: normalized to the interval [0...1]

- **Average Precision**: ranked precision metric that places emphasis on highly ranked correct predictions (hits)

Metrics for Rating Prediction

- Ground truth = ratings
- **Mean Absolute Error (MAE)**: computes the deviation between predicted ratings and actual ratings

- **Root Mean Square Error (RMSE)**: similar to MAE, but places more emphasis on larger deviation

Online Evaluation

Characteristics of methods

- Subject
- Research Method
- Setting

Evaluation Settings

Lab Studies

- Expressly created for the purpose of the study
- Extraneous variables can be controlled more easily by selecting study participants who should behave as they would in a real-world environment but doubts may exist about participants motivated by money, prizes or social pressure

Field Studies

- Conducted in a preexisting real world environment
- Users are intrinsically motivated to use a system

Experimental Design

- Hypotheses on personalized vs. non-personalized recommendation techniques and their potential to do something

Non-experimental Research

- **Quasi-experiments**: lack random assignments of units to different treatments
- **Non-experimental/observational research**: surveys/questionnaires, longitudinal research, case studies, focus group

Data

Types of Ratings

Explicit

- typical choices
- possibly multidimensional
- **main challenge**: users not always willing to rate many items

Implicit

- user action interpreted as rating
- easy to collect transparently without additional effort
- **main challenge**: action doesn't necessarily have the same meaning as a rating

Sparsity

- Data is sparse
- Natural datasets include historical interaction records of real users
- Sparsity can be measured $Sparsity = 1 - |R|/|I|.|U|$, R = ratings, I = Items, U = Users

Problems

- How many items in common are 2 users expected to have?
- **Cold start**: How to recommend new items? What to recommend to new users
  - Ask/force users to rate a set of items
  - In the beginning use method not based on rating
  - Default voting

More Algorithms

Memory-based approaches

- User-based CF
  - rating matrix is directly used to find neighbors/make predictions
  - does not scale for most real-world scenarios

Model-based approaches

- Based on an offline pre-processing or "model-learning" phase
- At run-time: only the learned model is used to make predictions
- Models are updated/re-trained periodically
- Matrix factorization techniques, statistics
- Association rule mining
- Probabilistic models (clustering models, Bayesian networks, probabilistic Latent Semantic Analysis)
- Various other machine learning approaches

Content-based recommendation

- **Content**: combination of attributes and (semi-)free text
  - Recommendation approach: related to NLP and document classification

Knowledge-based recommendations

- Users want to define their requirements explicitly
- Time span plays an important role
- Items with low number of available ratings
- **Constraint-based**: based on explicitly defined set of recommendation rules; fulfill recommendation rules
- **Case-based**: based on different types of similarity measures; retrieve items that are similar to specified requirements
- Both approaches are similar in their **conversational** recommendation process
  - users specify the requirements
  - systems try to identify solutions
  - if no solution can be found, users change requirements

Interaction: critiquing

- User may not know exactly what they are seeking but can specify why their current item is not satisfactory

Hybrid Recommender Systems

- Monolithic exploitation of different features
- Parallel
- Pipeline

Challenges

Explanation

2 parties involved:

- organization interested in convincing user
- user concerned about making the right choice(s)

Attacks

- (monetary) value of being in recommendation lists
- Attacks aim to:
  - push some items
  - sabotage other items
  - simply sabotage the system
  - manipulation the "internet opinion"

Research Questions in Ubiquitous Domains

- **Goals**: serendipitous recommendations vs. proximity
- role of contextual parameter
- modality of interaction for users "on the go"

Application domains

- M-Commerce
- Tourism and visitor guide
- Cultural heritage and museum guides
- Home computing and entertainment

[< Go back]