Rethinking Collaborative Filtering: A Practical Perspective on State-Of-The-Art Research Based on “Real-World” Insights and Challenges

Noam Koenigstein
RECOMMENDATIONS IN MICROSOFT STORE
The Xbox Marketplace
<table>
<thead>
<tr>
<th>Track</th>
<th>Artist</th>
<th>Album</th>
<th>Year</th>
<th>Genre</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vogue</td>
<td>Madonna</td>
<td>Celebration (Double Disc)</td>
<td>2009</td>
<td>Pop</td>
<td>5:16</td>
</tr>
<tr>
<td>Well, Well</td>
<td>Nelly Furtado</td>
<td>Whoa, Nelly!</td>
<td>2000</td>
<td>Pop</td>
<td>2:59</td>
</tr>
<tr>
<td>1,2,3 (Remix)</td>
<td>Gloria Estefan</td>
<td>Greatest Hits</td>
<td>1992</td>
<td>Pop</td>
<td>3:34</td>
</tr>
<tr>
<td>Viva Forever</td>
<td>Spice Girls</td>
<td>Spiceworld</td>
<td>1997</td>
<td>Rock</td>
<td>5:10</td>
</tr>
<tr>
<td>Man in the Mirror (2012 Rem)</td>
<td>Michael Jackson</td>
<td>Bad (Remastered)</td>
<td>1987</td>
<td>Pop</td>
<td>5:18</td>
</tr>
<tr>
<td>Hung Up</td>
<td>Madonna</td>
<td>Confessions On A Dance</td>
<td>2005</td>
<td>Pop</td>
<td>5:37</td>
</tr>
<tr>
<td>So What (Main Version)</td>
<td>P!nk</td>
<td>Funhouse</td>
<td>2009</td>
<td>Pop</td>
<td>3:35</td>
</tr>
</tbody>
</table>
Microsoft’s Web-Store

Students save 10% on new Surface Pro. SHOP NOW →

BUNDLE AND SAVE

New Surface Pro i5 128GB + Type Cover

★ ★ ★ ★ ★ (4)

Build your bundle
Free shipping, Free returns.

Description
• Save $150
• Intel Core i5, 128GB SSD, 4GB RAM
• Choice of Surface Type Cover (Burgundy, Cobalt Blue, or Platinum)
• A best-in-class laptop, with the versatility of a studio and tablet
• Powerful Intel Core processor
• The fastest startup and resume of any Surface Pro yet
• Up to 13.5 hours of battery life
• Surface Pen sold separately
A DECADE AGO...
A Decade Ago… The Netflix Prize

The goal: 10% improvement in RMSE over Netflix’s Cinematch

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_{\uparrow i} - y_{\downarrow i})^2}$$

It took tens of thousands of participants over 2 years…. 
The Problem with Ratings

• They do not exist!

• Missing items not at random
  Collaborative Filtering and the Missing at Random Assumption
  B. M. Marlin, R. S. Zemel, S. Roweis, M. Slaney

• Ratings are fuzzy and influenced by the order of items
  G. N. Yannakakis, J. Hallam

• Learning ratings is very different from personalization!
  Yahoo! Music Recommendations: Modeling Music Ratings with Temporal Dynamics and Taxonomy
  Gideon Dror, Noam Koenigstein and Yehuda Koren
IF NOT RMSE THEN WHAT?
Implicit Feedback and Ranking

- Collaborative Filtering for Implicit Feedback Datasets
  Y. Hu, Y. Koren, C. Volinsky

- Implicit-to-Explicit Ordinal Logistic Regression
  D. Parra, A. Karatzoglou, X. Amatriain, I. Yavuz

- BPR - Bayesian Personalized Ranking
  S. Rendle, C. Freudenthaler, Z. Gantner, and L. S. Thieme

- RankALS – Alternating Least Squares for Personalized Ranking
  G. Takacs, D. Tikk

- CLiMF – Reciprocal Rank Optimization
  Y. Shi, A. Karatzoglou, L. Baltrunas, M. Larson, N. Oliver, A. Hanjalic
ONE-CLASS COLLABORATIVE FILTERING
WITH RANDOM GRAPHS

Ulrich Paquet and Noam Koenigstein

*International World Wide Web Conference (WWW'13), May 2013, Rio de Janeiro, Brazil.*
Problem Formulation

Bipartite graph $\rightarrow$ We care about $\bar{p}(\text{link})$

$M \approx 10 - 100\text{M nodes}$

$N \approx 10 - 100\text{K nodes}$
The Hidden Graph

\[ G = \{ g_{\downarrow mn} \}, H = \{ h_{\downarrow mn} \} \]

\[ u \downarrow m \]

\[ g_{\downarrow mn} = 1 \quad h_{\downarrow mn} = 1 \]

\[ g_{\downarrow mn} = 0 \quad h_{\downarrow mn} = 1 \]

\[ G = \{ g_{\downarrow mn} \}, H = \{ h_{\downarrow mn} \} \]

edges \( g, h \in \{0,1\} \)

\[ u \uparrow T v \]

\[ pg = 1 \quad u, v, h = 1 = \sigma(u \uparrow T v) \]

\[ u \downarrow m \]

\[ v \downarrow n \]
BESIDES FEEDBACK: COLD START, META-DATA, HYBRID, CONTEXTUAL...
XBOX MOVIES RECOMMENDATIONS: VARIATIONAL BAYES MATRIX FACTORIZATION WITH EMBEDDED FEATURE SELECTION

Noam Koenigstein and Ulrich Paquet

ACM Conference on Recommender Systems (RecSys'13), October 2013, Hong Kong, China.
Harry Potter and the Philosopher's Stone

Categories:
- Plot
- Mood
- Audience
- Time Period

- Imaginary
- Wizards and Magicians
- Best Friends
- Exciting
- Humorous
- Danger
- Kids
- Teens
- Contemporary
- 21st Century
\[ p(f \downarrow 1, f \downarrow 2 \mid \alpha = 0.01, \beta = 0.01) \]
GROOVE RADIO: A BAYESIAN HIERARCHICAL MODEL FOR PERSONALIZED PLAYLIST GENERATION

Shay Ben-Elazar, Gal Lavee, Noam Koenigstein, Oren Barkan, Hilik Berezin, Ulrich Paquet, Tal Zaccai

*ACM Conference on Web Search and Data Mining (WSDM'17), Cambridge UK, February 2017.*
THE GAP BETWEEN COLLABORATIVE FILTERING RESEARCH AND REAL WORLD RECOMMENDATIONS
The Gap Between Collaborative Filtering and Real Recommenders

- Diversity vs. accuracy - tradeoff??
- Popularity vs. personalization
- Item fatigue / freshness – repeating items
- Serendipity – when and how much to “surprise” the user
- List Recommendations / page optimization
- Predicting the future vs. influencing the user
- Metrics and Evaluation
The Salesperson Analogy
BEYOND COLLABORATIVE FILTERING: THE LIST RECOMMENDATION PROBLEM

Oren Sar Shalom, Noam Koenigstein, Ulrich Paquet, Hastagiri P. Vanchinathan

*International World Wide Web Conference (WWW'16)*, April 2016, Montreal, Canada.
List Recommendations in Xbox 360

- Halo 4
- Call of Duty: Black Ops II
- FIFA Soccer 13
- Nike+ Kinect Training
- Grand Theft Auto V
- Forza Horizon 2
Conclusions

• There is still a gap between most CF models and the actual goal of recommender systems

• Learning individual user-item tuples or ranking preferences is problematic because:
  – Can’t handle the diversity vs. accuracy “tradeoff”
  – List recommendations / Page optimization

• Learning to predict future events from historical data is insufficient because:
  – Can’t handle balancing popularity and personalization
  – Freshness / Item Fatigue
  – Serendipity

• RL alone is not the ultimate solution because:
  – The abundance of implicit data
  – Representing the “taste space”

• Offline evaluation metrics are insufficient
  – They measure our ability to predict the future but not our ability to change it (influence the user)

• Bottom line: We still have a lot to work in the RecSys community!
Thank You!

We are looking for postdocs in Israel!!!

Interested?
Find me during the coffee break....