Recommender Systems in Finance: Methods and Business Value

Dietmar Jannach, University of Klagenfurt, Austria
dietmar.jannach@aau.at

Presented at the Workshop on Recommender Systems in Finance
Bergen, Norway, 2020
Recommender Systems (RS)

- A pervasive part of our daily online user experience
- One of the most widely used applications of machine learning / artificial intelligence
# Applications

<table>
<thead>
<tr>
<th>Category</th>
<th>Subcategories</th>
</tr>
</thead>
<tbody>
<tr>
<td>News</td>
<td>Apps</td>
</tr>
<tr>
<td>Books</td>
<td>Restaurants</td>
</tr>
<tr>
<td>Videos</td>
<td>Hotels</td>
</tr>
<tr>
<td>Music</td>
<td>Deals</td>
</tr>
<tr>
<td>Games</td>
<td>Partners</td>
</tr>
<tr>
<td>Shopping goods</td>
<td>...</td>
</tr>
<tr>
<td>Friends</td>
<td>Cigars</td>
</tr>
<tr>
<td>Groups</td>
<td>Software code, code reviewers</td>
</tr>
<tr>
<td>Jobs</td>
<td>...</td>
</tr>
</tbody>
</table>
Applications

• Where’s Finance?
• Recommendations obviously play a role in Finance
Recommender Systems in Finance

- Some activity happening in academia

- But far from being mainstream
Outline

- Understand the potential values of recommender systems
- Learn how we they are built
- [Understand the needs in the financial domain]
Who benefits?

• Why should we use recommender systems?
  – Recommenders can have value both for consumers and the providers of the recommendations
  – There can be even more stakeholders:
    • e.g.: consumers, trader, product provider

Potential value for the consumer

• Examples:
  – Help users find objects that match their long-term preferences (information filtering)
  – Help users explore the item space and improve decision making
  – Make contextual recommendations, e.g.,
    • Show alternatives
    • Show accessories
  – Remind users of what they liked in the past
  – Actively notify consumers of relevant content
Potential value for the provider

• Examples:
  – Change **user behavior** in desired directions
  – Create additional **demand**
  – Increase (short term) **business success**
  – Enable item “**discoverability**”
  – Increase activity on the site and **user engagement**
  – Provide a valuable **add-on service**
  – **Learn more** about the customers
Multi-stakeholder considerations

• When goals are fully aligned
  – Better recommendations can lead to more satisfied, returning customers who find what they need

• When there can be a goal conflict
  – Not all recommendable items may have the same business value
  – From a business perspective, it might be better to recommend items with a higher sales margin
    • As long as the recommendations are still reasonable
What’s the business value?

• Typical quotes about value

“35% of Amazon.com’s revenue is generated by its recommendation engine.”

“We think the combined effect of personalization and recommendations save us more than $1B per year.”

“Netflix says 80 percent of watched content is based on algorithmic recommendations”

Measuring the business value

• Measuring the business value can be difficult
  – What does it tell us that 80% of the watched content comes from the recommendations?
  – Where do the said savings come from?
• The used measures often largely depend on
  – The business model of the provider
  – The intended effects of the recommendations
  – Assumptions about consumer value
What is often measured?

- Considering both the **impact** and **value** perspective

---

Click-Through Rates

• Measures how many clicks are garnered by recommendations
  – Popular in the news recommendation domain
    • Google News: 38% more clicks compared to popularity-based recommendations
    • Forbes: 37% improvement through better algorithm compared to time-decayed popularity based method
    • swissinfo.ch: Similar improvements when considering only short-term navigation behavior
  – YouTube: Almost 200% improvement through co-visitation method (compared to popular recommendations)
Adoption and Conversion Rates

- CTR usually not the ultimate measure
  - Cannot know if users actually liked/purchased what they clicked on (consider also: click bait)

- Therefore
  - Various, domain-specific adoption measures common

- YouTube, Netflix: “Long CTR”/ “Take rate”
  - only count click if certain amount of video was watched
Adoption and Conversion Rates

• Alternatives when items cannot be viewed/read:

  • eBay:
    – “purchase-through-rate”, “bid-through-rate”

  • Other:
    – LinkedIn: Contact with employer made
    – Paper recommendation: “link-through”, “cite-through”
    – E-Commerce marketplace: “click-outs”
    – Online dating: “open communications”, “positive contacts per user”
Sales and Revenue

• CTR and adoption measures are good indicators of relevant recommendations
• However:
  – Often unclear how this translates into business value
  – Users might have bought an item anyway
  – Substantial increases might be not relevant for business when starting from a very low basis
• In addition:
  – Problem of measuring effects with flat-rate subscription models (e.g., Netflix).
Sales and Revenue

• Only a few studies, some with limitations
  – Video-on-demand study: 15% sales increase after introduction (no A/B test, could be novelty effect)
  – DVD retailer study:
    • 35% lift in sales when using purchased-based recommendation method compared to “no recommendations”
    • Almost no effects when recommendations were based on view statistics
    • Choice of algorithm matters a lot
Sales and Revenue

- **e-grocery studies:**
  - 1.8% direct increase in sales in one study
  - 0.3% direct effects in another study
  - However:
    - Up to 26% indirect effects, e.g., where customers were pointed to other categories in the store
    - “Inspirational” effect also observed in music recommendation in our own work

- **eBay:**
  - 6% increase for similar item recommendations through largely improved algorithm
  - (500% increase in other study for specific area)
Sales and Revenue

- **Book store study:**
  - 28% increase with recommender compared with “no recommender”; could be seasonal effects
  - Drop of 17% after removing the recommender

- **Mobile games (own study)**
  - 3.6% more purchases through best recommender
  - More is possible
Effects on Sales Distributions

• Goal is maybe not to sell more but different items

• Influence sales behavior of customers
  – stimulate cross-sales
  – sell off on-stock items
  – promote items with higher margin
  – long-tail recommendations
Effects on Sales Distributions

- Premium cigars study:
  - Interactive advisory system installed
  - Measurable shift in terms of what is sold
    - e.g., due to better-informed customers
Effects on Sales Distributions

• Netflix:
  – Measure the “effective catalog size”, i.e., how many items are actually (frequently) viewed
  – Recommenders lead users away from blockbusters
    • Could also be beneficial in terms of license costs

• Online retailer study:
  – Comparison of different algorithms on sales diversity
  – Outcomes
    • Recommenders tend to decrease the overall diversity
    • Might increase diversity at individual level though
User Behavior and Engagement

• Assumption:
  – Higher engagement leads to higher re-subscription rates (e.g., at Spotify)

• News domain studies:
  – 2.5 times longer sessions, more sessions when there is a recommender

• Music domain study:
  – Up to 50% more user activity

• LinkedIn:
  – More clicks on job profiles after recommender introduced
Discussion & challenges

• General
  – No doubt about huge business potential of RS

• Challenges, e.g.:
  – Direct financial impact sometimes difficult to measure (e.g., Netflix)
  – User activity and clicks might not be indicative of consumer value and business success
  – Long-term effects often difficult to measure
  – Know what to measure
What to measure?

• Some underlying questions:
  – What is the intended purpose of the system?
  – What kind of value should it create?
  – How can we assess (and balance) the value for the different stakeholders?

A conceptual framework

- Should help to decide what and how to measure (both in academia and industry)
- Layered structure – strategic to operational
- Considers two viewpoints

<table>
<thead>
<tr>
<th>Overarching goal of the system, strategic value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommendation purpose / Intended utility</td>
</tr>
<tr>
<td>System (algorithm) task</td>
</tr>
<tr>
<td>Computational metrics</td>
</tr>
</tbody>
</table>
Framework overview

<table>
<thead>
<tr>
<th>Strategic Perspective</th>
<th>Consumer's Viewpoint</th>
<th>Provider's Viewpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overarching Goal</td>
<td>“Personal Utility”: Happiness, Satisfaction, Knowledge, ...</td>
<td>“Organizational Utility”: Profit, Revenue, Growth, ...</td>
</tr>
</tbody>
</table>
| Recommendation Purpose | • Help users find objects that match the user’s long-term preferences  
• Show alternatives  
• Help users explore or understand the item space  
• ... | • Change user behavior in desired directions  
• Create additional demand  
• Increase activity on the site  
• ... |
| System Task           | • Annotate in context (i.e., estimate preference of a given item)  
• Find good items  
• Create diverse set of alternatives  
• Find suitable accessories  
• Retrieve novel but relevant items  
• ... |
<p>| Operational Perspective | Computational Metric | Predictive accuracy (e.g., RMSE, MAE), classification accuracy (e.g., precision, recall, AUC), ranking and top-n accuracy (e.g., rank correlation, MRR, NDCG, etc.), item “discoverability” (diversity, novelty, or serendipity measures), recommendation biases (e.g., concentration or popularity biases) and blockbuster effects, survey-based user satisfaction scores, business- and domain-specific measures (e.g., conversion rates or click-through-rates), ... |</p>
<table>
<thead>
<tr>
<th>Strategic Perspective</th>
<th>Consumer’s Viewpoint</th>
<th>Provider’s Viewpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overarching Goal</strong></td>
<td>&quot;Personal Utility&quot;: Happiness, Satisfaction, Knowledge, Entertainment, Benefit</td>
<td>&quot;Organizational Utility&quot;: Profit, Revenue, Return on Investment, Growth, Customer Retention</td>
</tr>
</tbody>
</table>
| **Recommendation Purpose** | • Help users find objects that match the user's long-term preferences  
• Show alternatives  
• Help users explore or understand the item space, ... | • Change user behavior in desired directions  
• Create additional demand  
• Help users discover new artists, directors, genres  
• Increase activity on the site  
• ... |
| **System Task** | • Annotate in context (i.e., estimate preference of a given item)  
• **Find good items**  
• Create diverse set of alternatives  
• Find mix of familiar and relevant unknown items  
• Find suitable accessories  
• ... | Predictive accuracy (e.g., RMSE, MAE), classification accuracy (e.g., Precision, Recall, AUC), ranking and top-n accuracy (e.g., rank correlation, MRR, NDCG, etc.), item discoverability (diversity, novelty, or serendipity measures), recommendation biases (e.g., concentration or popularity biases) and blockbuster effects, survey-based user satisfaction scores, business- and domain-specific measures (e.g., conversion rates or click-through-rates), ... |
| **Computational Metric** | | |
### Consumer’s Viewpoint

**Overarching Goal**
- "Personal Utility": Happiness, **Satisfaction**, Knowledge, Entertainment, Benefit

**Recommendation Purpose**
- Help users find objects that match the user’s long-term preferences
  - Show alternatives
  - Help users explore or understand the item space, ...

### Provider’s Viewpoint

**Overarching Goal**
- "Organizational Utility": Profit, Revenue, Return on Investment, Growth, Customer Retention

**Recommendation Purpose**
- Change user behavior in desired directions
- Create additional demand
- Help users discover new artists, directors, genres
- Increase activity on the site
- ...

### Operational Perspective

**System Task**
- Annotate in context (i.e., estimate preference of a given item)
- **Find good items**
- Create diverse set of alternatives
- Find mix of familiar and relevant unknown items
- Find suitable accessories
- ...

**Computational Metric**
- Predictive accuracy (e.g., RMSE, MAE), classification accuracy (e.g., **Precision, Recall, AUC**), ranking and top-n accuracy (e.g., rank correlation, MRR, NDCG, etc.), item discoverability (diversity, novelty, or serendipity measures), recommendation biases (e.g., concentration or popularity biases) and blockbuster effects, survey-based user satisfaction scores, business- and domain-specific measures (e.g., conversion rates or click-through-rates), ...

<table>
<thead>
<tr>
<th><strong>Strategic Perspective</strong></th>
<th><strong>Consumer’s Viewpoint</strong></th>
<th><strong>Provider’s Viewpoint</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overarching Goal</strong></td>
<td>&quot;Personal Utility&quot;: Happiness, Satisfaction, Knowledge, Entertainment, Benefit</td>
<td>&quot;Organizational Utility&quot;: Profit, Revenue, Return on Investment, Growth, Customer Retention</td>
</tr>
</tbody>
</table>
| **Recommendation Purpose** | • Help users find objects that match the user’s long-term preferences  
• Show alternatives  
• Help users explore or understand the item space, ... | • Change user behavior in desired directions  
• Create additional demand  
• Help users discover new artists, directors, genres  
• Increase activity on the site  
• ... |

<table>
<thead>
<tr>
<th><strong>Operational Perspective</strong></th>
<th><strong>System Task</strong></th>
<th><strong>Computational Metric</strong></th>
</tr>
</thead>
</table>
| **Overarching Goal**        | • Annotate in context (i.e., estimate preference of a given item)  
• **Find good items**  
• Create diverse set of alternatives  
• Find mix of familiar and relevant unknown items  
• Find suitable accessories  
• ... | Predictive accuracy (e.g., RMSE, MAE), classification accuracy (e.g., Precision, Recall, AUC), ranking and top-n accuracy (e.g., rank correlation, MRR, NDCG, etc.), item discoverability (diversity, novelty, or serendipity measures), recommendation biases (e.g., concentration or popularity biases) and blockbuster effects, **survey-based user satisfaction scores**, business- and domain-specific measures (e.g., conversion rates or click-through-rates), ... |
<table>
<thead>
<tr>
<th>Strategic Perspective</th>
<th>Consumer’s Viewpoint</th>
<th>Provider’s Viewpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overarching Goal</strong></td>
<td>&quot;Personal Utility&quot;: Happiness, Satisfaction, Knowledge, Entertainment, Benefit</td>
<td>&quot;Organizational Utility&quot;: Profit, Revenue, Return on Investment, Growth, Customer Retention</td>
</tr>
</tbody>
</table>
| **Recommendation Purpose** | • Help users find objects that match the user's long-term preferences  
• Show alternatives  
• Help users explore or understand the item space, ... | • Change user behavior in desired directions  
• Create additional demand  
• Help users discover new artists, directors, genres  
• Increase activity on the site  
• ... |
| **System Task** | • Annotate in context (i.e., estimate preference of a given item)  
• Find good items  
• Create diverse set of alternatives  
• Find mix of familiar and relevant unknown items  
• Find suitable accessories  
• ... | Predictive accuracy (e.g., RMSE, MAE), classification accuracy (e.g., Precision, Recall, AUC), ranking and top-n accuracy (e.g., rank correlation, MRR, NDCG, etc.), item discoverability (diversity, novelty, or serendipity measures), recommendation biases (e.g., concentration or popularity biases) and blockbuster effects, survey-based user satisfaction scores, business- and domain-specific measures (e.g., conversion rates or click-through-rates), ... |
<p>| <strong>Operational Perspective</strong> | <strong>Computational Metric</strong> |</p>
<table>
<thead>
<tr>
<th>Overarching Goal</th>
<th>&quot;Personal Utility&quot;: Happiness, Satisfaction, Knowledge, Entertainment, Benefit</th>
<th>&quot;Organizational Utility&quot;: Profit, Revenue, Return on Investment, Growth, <strong>Customer Retention</strong></th>
</tr>
</thead>
</table>
| Recommendation Purpose | • Help users find objects that match the user's long-term preferences  
• Show alternatives  
• Help users explore or understand the item space, ... | • Change user behavior in desired directions  
• Create additional demand  
• Help users discover new artists, directors, genres  
• Increase activity on the site  
• ... |
| System Task | • Annotate in context (i.e., estimate preference of a given item)  
• Find good items  
• Create diverse set of alternatives  
• Find mix of familiar and relevant unknown items  
• Find suitable accessories  
• ... | |
<p>| Computational Metric | Predictive accuracy (e.g., RMSE, MAE), classification accuracy (e.g., Precision, Recall, AUC), ranking and top-n accuracy (e.g., rank correlation, MRR, NDCG, etc.), item discoverability (diversity, novelty, or serendipity measures), recommendation biases (e.g., concentration or popularity biases) and blockbuster effects, survey-based user satisfaction scores, business- and domain-specific measures (e.g., conversion rates or click-through-rates), ... | |</p>
<table>
<thead>
<tr>
<th>Strategic Perspective</th>
<th>Consumer’s Viewpoint</th>
<th>Provider’s Viewpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overarching Goal</strong></td>
<td>&quot;Personal Utility&quot;: Happiness, Satisfaction, Knowledge, Entertainment, Benefit</td>
<td>&quot;Organizational Utility&quot;: Profit, Revenue, Return on Investment, Growth, <strong>Customer Retention</strong></td>
</tr>
</tbody>
</table>
| **Recommendation Purpose** | • Help users find objects that match the user's long-term preferences  
• Show alternatives  
• Help users explore or understand the item space, ... | • Change user behavior in desired directions  
• Create additional demand  
• **Help users discover new artists, directors, genres**  
• Increase activity on the site  
• ... |
| **System Task** | • Annotate in context (i.e., estimate preference of a given item)  
• Find good items  
• Create diverse set of alternatives  
• Find mix of familiar and relevant unknown items  
• Find suitable accessories  
• ... | Predictive accuracy (e.g., RMSE, MAE), classification accuracy (e.g., Precision, Recall, AUC), ranking and top-n accuracy (e.g., rank correlation, MRR, NDCG, etc.), item discoverability (diversity, novelty, or serendipity measures), recommendation biases (e.g., concentration or popularity biases) and blockbuster effects, survey-based user satisfaction scores, business- and domain-specific measures (e.g., conversion rates or click-through-rates), ... |
<p>| <strong>Operational Perspective</strong> | <strong>Computational Metric</strong> |</p>
<table>
<thead>
<tr>
<th>Strategic Perspective</th>
<th>Consumer’s Viewpoint</th>
<th>Provider’s Viewpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overarching Goal</strong></td>
<td>&quot;Personal Utility&quot;: Happiness, Satisfaction, Knowledge, Entertainment, Benefit</td>
<td>&quot;Organizational Utility&quot;: Profit, Revenue, Return on Investment, Growth, <strong>Customer Retention</strong></td>
</tr>
</tbody>
</table>
| **Recommendation Purpose** | • Help users find objects that match the user's long-term preferences  
• Show alternatives  
• Help users explore or understand the item space, ... | • Change user behavior in desired directions  
• Create additional demand  
• **Help users discover new artists, directors, genres**  
• Increase activity on the site  
• . . . |
| **System Task**        | • Annotate in context (i.e., estimate preference of a given item)  
• Find good items  
• Create diverse set of alternatives  
• **Find mix of familiar and relevant unknown items**  
• Find suitable accessories  
• ... | |
<p>| <strong>Operational Perspective</strong> | Predictive accuracy (e.g., RMSE, MAE), classification accuracy (e.g., Precision, Recall, AUC), ranking and top-n accuracy (e.g., rank correlation, MRR, NDCG, etc.), item discoverability (diversity, novelty, or serendipity measures), recommendation biases (e.g., concentration or popularity biases) and blockbuster effects, survey-based user satisfaction scores, business- and domain-specific measures (e.g., conversion rates or click-through-rates), . . . |</p>
<table>
<thead>
<tr>
<th>Overarching Goal</th>
<th>Consumer’s Viewpoint</th>
<th>Provider’s Viewpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Personal Utility&quot;: Happiness, Satisfaction, Knowledge, Entertainment, Benefit</td>
<td></td>
<td>&quot;Organizational Utility&quot;: Profit, Revenue, Return on Investment, Growth, <strong>Customer Retention</strong></td>
</tr>
</tbody>
</table>

### Recommendation Purpose
- Help users find objects that match the user’s long-term preferences
- Show alternatives
- Help users explore or understand the item space, ...
- Change user behavior in desired directions
- Create additional demand
- **Help users discover new artists, directors, genres**
- Increase activity on the site
- ...

### System Task
- Annotate in context (i.e., estimate preference of a given item)
- Find good items
- Create diverse set of alternatives
- **Find mix of familiar and relevant unknown items**
- Find suitable accessories
- ...

### Operational Perspective

#### Computational Metric

- **Predictive accuracy** (e.g., RMSE, MAE), classification accuracy (e.g., Precision, Recall, AUC), ranking and top-n accuracy (e.g., rank correlation, MRR, NDCG, etc.), item discoverability (diversity, **novelty, or serendipity measures**), recommendation biases (e.g., concentration or popularity biases) and blockbuster effects, survey-based user satisfaction scores, business- and domain-specific measures (e.g., conversion rates or click-through-rates), ...
Summary of value considerations

• Demonstrated business value of recommenders in many domains

• Size of impact however depends on many factors like baselines, domain specifics etc.

• Measuring impact is generally not trivial
  – Choice of the evaluation measure matters a lot
  – CTR can be misleading

• “Metric-Task-Purpose-Fit” to be considered
Methods
Outline

• Content-based Filtering
• Collaborative Filtering
• Hybrid Systems
• Knowledge-based Systems

• Interactive Recommendation
Recommendation Principles

Recommendations are usually personalized.
Content-based Filtering

Content-based:
"Show me more of the same what I've liked"
Because you liked Kaskade & Felix Cartal - Fakin' It (feat. Ofelia K)

- Cazzette 'Blue Skies Ahead...
  PRMD Music

- Neon Owl Radio 12: TELYK...
  Neon Owl

- City Of Angels
  BYNON

- Breathe
  EMBRZ
Collaborative Filtering

Collaborative: "Tell me what's popular among my peers"
Customers who bought ...
Hybrid Recommendation Approach

Hybrid:
Combinations of various inputs and/or composition of different mechanism
Knowledge-based Systems

Knowledge-based: "Tell me what fits based on my needs"
Knowledge-based Systems

• Hybrid collaborative and content-based systems are highly successful in practice
  – And do not require knowledge engineering
• Why do we need knowledge-based systems?
  – Often low number of past interactions (purchases)
  – Often explicit customer requirements are relevant
  – Longer time spans between purchases
An interactive travel recommender
An interactive travel recommender
An interactive travel recommender

http://www.configworks-gmbh.online.de - VIBE - the virtual adviser for the Warmbad-Villach spa resort

You’re bound to ask yourself why I recommended the following. I’ll be happy to explain...

My arguments specially for you.

I am happy to have found autumn packages for you, as you wished. If you want more suggestions for a specific date, you’ll have to use the detailed advice option (more questions).

We have a whole range at the Warmbad-Villach spa resort to suit your request Leisure and activities programme & Long walks. Ask about them.

Our comprehensive supporting programme of cultural events (Corinthian Summer Music Festival, Villach Carnival), exhibitions at the Warmbad culture club, Jazz Over Villach, etc) all year round and attractions in the vicinity will round off your stay at the

Do you want to feel fit and healthy? Our sports and activities programme respond to your wishes.
A financial advisory system

- Built for an Austrian bank in mid-2000s
- Provides multi-step interactive needs acquisition
  - personalized according to preferences
- Generates tailored set of recommendations
- Is able to explain recommendations
- Points consumers to inconsistencies with respect to their expectations
Financial advisor demo (German)
Discussion

- Successful application in practice (own startup)
  - Built more solutions, e.g. for Hungarian bank or Austrian insurance company
- Mimics behavior of experienced sales advisor
  - Automated documentation of advisory process
- Solutions based on expert knowledge
  - Comprehensive tooling, automated generation of application
  - No learning component
Discussion

• In 2020:
  – Chatbots have become popular in 2016
  – Enormous advances in natural language technology

• Challenges:
  – Limitations of pure learning approaches for conversational recommendation
  – Certain amounts of knowledge engineering required
    • Predictability of recommendations is an issue
Examples of other applications

- Content / Case-based Recommendation
  - Look up similar past customers
Case-based recommendation

Fig. 2. Our case-based recommendation pipeline.
Examples of other applications

- A number of portfolio advisory approaches
  - Mostly knowledge-intensive
Examples of other applications

- Equity funds selection
  - Collaborative content-based hybrid
  - Multi-criteria decision making
Discussion

• Undisputed success of recommender systems in many domains

• Ways forward in the financial domain
  – Financial recommendations Amazon style?
    • Customers who bought ...
  – Conversational recommendation?
  – Specific applications?
    • e.g., portfolio or investment recommendations
  – Mostly data analytics?
    • forecasting – also a sort of decision-support and recommendation
• Thank you for your attention
• dietmar.jannach@aau.at

• https://tinyurl.com/finrec2020