

# 4-4: Explaining Recommendations

# Learning Objectives

- To understand the variety of ways that explanations are made in recommender systems, and how users react to them.
- To understand how explanations can fit into a recommender system
- To recognize pitfalls in explanation – and particularly the difference between explanation and persuasion

# Key Reference

- Explaining Collaborative Filtering Recommendations by Herlocker, Konstan, and Riedl (Proc. CSCW 2000)
  - Several dozen more recent papers look at specific aspects – transparency, trust, computing efficient explanations, and more ...

# What is an Explanation?

- Additional data to help users understand a specific recommendation
  - Separate from an explanation of how the system works as a whole
  - Some explanations are confidence scores
- Often a glimpse inside the computation
  - More detail than a simple prediction score ...
    - Predictions don't distinguish normal from weird distributions
- Sometimes tied to ability to edit profile to improve recommendations ...

# Let's take a look at Herlocker ...

- Key lessons
  - Simplicity is key – users didn't like “correlation,” statistical terms, or being overwhelmed with data
  - Simple visualizations work well – histogram, table
  - Supporting information valued – historical success, attribute-linked data, associations
- One key mistake
  - We didn't actually measure explanation effectiveness, but rather persuasiveness

# Explanations in Practice

- We saw Amazon.com earlier
- “Why?” is a compelling anchor
- Tell me more provides user control
- Be careful, though – overwhelming the user is rarely a good strategy
- Explanations not only for user-user CF
- Mostly open area: explaining “lists” vs. items

# Take-Away

- You've now seen several types of explanation:
  - More data on the prediction itself
  - Other data about the item
  - Simplified presentations of relevant data
  - Past performance data
  - Supporting statistics
- You now have some experience with which explanations work with ordinary users ...

# Moving Forward

- Next
  - Interviews related to trust, reputation, influence of bad ratings, alternative ways to focus on trusted sources or limit unknown ones
  - Assignments related to user-user recommendation
  - Next module focuses on evaluation – will also provide new metrics that could help with explanation!



# 4-4: Explaining Recommendations