

# Explainable Recommender System

July 11, 2018

## 1 Introduction

Many recommender systems functioned as black boxes, providing no transparency into the working of the recommendation process, nor offering any additional information to accompany the recommendations beyond the recommendations themselves.

**Definition of Explanation:** Explanations can benefit recommender systems on a number of aspects. Thus, based on their goals, we can distinguish between different explanations [5].

**Transparency explanation** explains how the system works. Transparency is usually evaluated in terms of its effect on actual and perceived understanding of how the system works, e.g. in [1] during a period of three weeks, 82 subjects are drawn to participate in a procedure of questionnaire, interview and follow-up questionnaire to give a quantitative analysis of transparency and acceptance.

**Scrutability explanation** allows users to tell the system it is wrong. The evaluations include metrics such as task correctness, information exposure (i.e. if users could express an understanding of what information was used to make recommendations for them)

**Trust explanation** increases users' confidence in the system. Questionnaires can be used to determine the degree of trust a user places in a system. Implicit measures such as loyalty are also commonly used.

**Effectiveness explanation** helps users make better decisions. Effectiveness can be calculated as perceived effectiveness, i.e. "this explanation helps me determine how well I will like this movie", or the absence of a difference between the liking of the recommended item prior to, and after consumption, or cross-system analysis based on how well the recommendations suit to personal tastes. Effectiveness could be also measured globally, i.e. with a marketing metric.

**Persuasiveness explanation** convince users to try or buy. Persuasiveness can be evaluated by user studies (i.e. questionnaire or 7-point Likert rating) of different explanation interfaces, or in different forms of "conversion rates", e.g. ratings before and after explanation, how much likely users actually buy items compared to users in a system without an explanation facility. It is worth noting that in [6], persuasiveness is closely related to social factors. A more credible source is more persuasive. A source whose intent is to persuade is

perceived as less trustworthy. Computers that convey similar personality types are more persuasive. It is thus important to integrate social cues to create more persuasive systems.

**Efficiency explanation** helps users make decisions faster. Efficiency is often evaluated in conversational recommender systems. In systems with which users rarely interact, we could evaluate efficiency by the number of activations of repair actions.

**Satisfaction explanation** increases the ease of use or enjoyment. When measuring satisfaction, one can directly ask users to finish a questionnaire. However, it is important to differentiate between satisfaction with the recommendation process and the recommended products.

## 2 Related Work

### 2.1 Explanation Sources

One major source of the explanations comes from collaborative filtering. For example, the typical "people also bought" scheme in Amazon and most other commercial systems. A number of previous research focuses on extracting structured information from social tags [4] and opinionated reviews [7, 2].

### 2.2 Explanation Styles

The most common explanation is expressed in a template, i.e. "customers who bought this item also bought ". It is also possible to place the extracted opinion in each aspect to form a structured overview, i.e. "we also recommend the following products because they are cheaper and lighter but have lower processor speed." Some systems use content similarity to organize the explanation. For example, Pandora highlights musical properties such as tempo and tonality and offers explanation such as "based on what you've told us so far, we're playing this track because it features a leisurely tempo and major tonality."

A number of work adopt visualization techniques, e.g. the ranking of different factors to provide the explanation. This category also includes visualizing associations via a knowledge graph.

### 2.3 Explanation and Recommendation

The underlying algorithm may influence the types of explanations that can be constructed. Collaborative filtering based explanations are typical examples. However, it is also possible that the explanations selected by the system developer do not reflect the underlying algorithm, especially when the recommendation algorithms are computationally complex. One possible solution is to embed the textual representation with the recommendation algorithm. For example, [3] aligns user/item latent factors in topical distributions. Topic modeling has been extensively used in this area [7, 2] due to its flexibility in combining with factor models .

## References

- [1] H. Cramer, V. Evers, S. Ramlal, M. Van Someren, L. Rutledge, N. Stash, L. Aroyo, and B. Wielinga. The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling and User-Adapted Interaction*, 18(5):455, 2008.
- [2] X. He, T. Chen, M.-Y. Kan, and X. Chen. Trirank: Review-aware explainable recommendation by modeling aspects. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, pages 1661–1670. ACM, 2015.
- [3] J. McAuley and J. Leskovec. Hidden factors and hidden topics: understanding rating dimensions with review text. In *Proceedings of the 7th ACM conference on Recommender systems*, pages 165–172. ACM, 2013.
- [4] A. Sharma and D. Cosley. Do social explanations work?: studying and modeling the effects of social explanations in recommender systems. In *Proceedings of the 22nd international conference on World Wide Web*, pages 1133–1144. ACM, 2013.
- [5] N. Tintarev and J. Masthoff. Designing and evaluating explanations for recommender systems. In *Recommender systems handbook*, pages 479–510. Springer, 2011.
- [6] K.-H. Yoo and U. Gretzel. Creating more credible and persuasive recommender systems: The influence of source characteristics on recommender system evaluations. In *Recommender systems handbook*, pages 455–477. Springer, 2011.
- [7] Y. Zhang, G. Lai, M. Zhang, Y. Zhang, Y. Liu, and S. Ma. Explicit factor models for explainable recommendation based on phrase-level sentiment analysis. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*, pages 83–92. ACM, 2014.