Dirichlet-Luce Choice Model for Online Learning from User Interactions

Invited Talk

Gökhan Çapan
İlker Gündoğdu
Ali Caner Türkmen
Ali Taylan Cemgil
Boğaziçi University
Istanbul, Turkey
gokhan.capan@boun.edu.tr

ABSTRACT

In modern digital applications, personalized user experience is shaped interactively. In a typical scenario, a recommender system estimates user preferences based on their previous choices. However, these choices are from only a subset of available items selected by the system in line with the user's preference estimates. In other words, users and the system interact in a feedback loop: the system learns to make future recommendations based on user choices from the alternatives it previously recommended.

In this talk, we describe a Bayesian choice model, the Dirichlet-Luce model, where we assume that choice observations comply with Luce's choice axiom, i.e., the users choose (or tend to choose) from a subset of all items recommended to them. Furthermore, the model is built on a generalization of the Dirichlet distribution as a prior probability distribution over user preferences (the joint distribution of choice probabilities), conjugate to the likelihood the choice observations lead to. We illustrate that the model achieves efficient inference of user preferences, based on the observation that the number of distinct presentations (subsets of the set of all items) that suffices for a 'good' preference estimate scales with the number of all items.

The Bayesian construction of the Dirichlet-Luce model leads to a bandit algorithm—based on Thompson sampling—for online learning to recommend. The algorithm achieves low-regret measured in terms of the inherent attractiveness of the items included in the recommendations, compared to several dueling bandits algorithms, a combinatorial bandit algorithm with relative feedback, and a state-of-the-art online learning-to-rank algorithm. Dirichlet-Luce model ensures independence of unexplored items. That is, posterior probabilities of preferences to the items that were never shown to the users (or are newly introduced to the system) stay invariant independently of other choices made. The combination of the Dirichlet-Luce model and the proposed bandit algorithm also eliminates some biases that recommender systems might be prone to; where the system overestimates user preference to promoted or initially preferred items due to overexposure or underestimates preference to underrepresented items in the recommendations. As a result, we believe the model has a potential to be reused as a fundamental building block for recommender systems.