Explaining Reviews and Ratings with PACO: Poisson Additive Co-Clustering

Chao-Yuan Wu∗
Department of Computer Science
University of Texas at Austin, Austin, TX
cywu@cs.utexas.edu

Alex Beutel∗
Computer Science Department
Carnegie Mellon University, Pittsburgh, PA
abeutel@cs.cmu.edu

Amr Ahmed
Strategic Technologies
Google, Mountain View, CA
amra@google.com

ABSTRACT
Understanding a user’s motivations provides valuable information beyond the ability to recommend items. Quite often this can be accomplished by perusing both ratings and review texts. Unfortunately matrix factorization approaches to recommendation result in large, complex models that are difficult to interpret. In this paper, we attack this problem through succinct additive co-clustering on both ratings and reviews. Our model yields accurate and interpretable recommendations.

Keywords
Co-clustering; recommendation systems; joint modeling

1. INTRODUCTION
Recommender systems often aim to generate suggestions while simultaneously explaining why a certain recommendation was made. We address this problem by extending ACCAMS [1] to include a novel additive language description in the form of a sum of Poisson distributions. This allows us to use backfitting for documents rather than just in a regression setting, and enables new applications.

With this approach we make a number of contributions:
• We design an additive co-clustering model, PACO, that can sum over both Gaussian and Poisson distributions. This allows us to use backfitting for documents rather than just in a regression setting, and enables new applications.
• We describe an efficient technique for sampling from a sum of Gaussian and Poisson random variables.
• We give empirical evidence across multiple datasets that PACO predicts ratings better than HFT [3] and JMARS [2]. Additionally, our method predicts reviews better than HFT, and achieves nearly as high quality prediction as JMARS, while being far faster and simpler. As seen in Figure 1, PACO outperforms both models in jointly predicting ratings and reviews.

∗These authors contributed equally.

2. ADDITIVE CO-CLUSTERING MODEL
In this section we give a high-level description on how to extend ACCAMS to jointly model ratings and reviews. See [1] for background and the complete paper [4] for details. We consider modeling a set of observed entries (u, m), each of which is comprised of a rating and a review that user u gives to item m. In the generative model of ACCAMS, each block in a co-clustering generates a Gaussian-distributed rating, a sum of which across co-clusterings gives the final rating. In PACO, each block further emits a Poisson-distributed word count, a sum of which across co-clusterings gives the final count n_{u,m,x} of word x for review (u, m), i.e.

\[ n_{u,m,x} \sim \text{Poi}(\lambda_{u,m,x}) \quad \text{and} \quad \mu_{x}^{(i)} \sim \text{Gamma}(\alpha, \beta) \]  

where

\[ \lambda_{u,m} = \mu^{(i)} + \mu^{(m)} \]

\[ = \sum_{i=1}^{S} \mu_{x}^{(i)} \cdot d_{i,m} + \mu_{u}^{(u)} \cdot e_{u,m} + \mu_{m}^{(m)} \cdot d_{m}^{(m)} \]  

and c_u and d_m are cluster assignments for (u, m). Here we further extend each co-clustering to have a user-clustering-specific language model \( \mu_{u}^{(u)} \) and an item-clustering-specific language model \( \mu_{m}^{(m)} \). In addition, we add a global item language model, \( \mu^{(m)} \), and a global background language model, \( \mu^{(0)} \). The text of the review is modeled as a combination of these Poisson language models.

2.1 Inference
We offer an efficient Gibbs sampling procedure to learn the PACO model. The collapsed Gibbs sampler for Gaussian distributions is described in [1]. In our complete paper [4], we give the precise equations for sampling a given \( \mu_{u,m,x}^{(i)} \) conditioned on all other language models in PACO. The key idea is a novel algorithm that parametrizes sampling from the sum of Poisson distributions as an efficient sampling from a multinomial distribution.

\[ \{\hat{n}_{u,m,x}\} \sim \text{Multi} \left( \frac{\{\mu_{u,m,x}\}}{\lambda_{u,m,x}}, n_{u,m,x} \right) \]  

This allows the model to attribute a given word of the review to each language model (user, movie or cluster).
Figure 1: Negative log likelihood. PACO better jointly predicts ratings and reviews than state-of-the-art JMARS \cite{diao2014jointly} and HFT \cite{mcauley2013hidden} on Amazon Fine Food, Yelp and RateBeer datasets. The joint predictive power is captured by the normalized negative log likelihood. Lower is better. (d) shows detailed results on IMDb dataset. More comparisons are given in \cite{wu2015explaining}.

<table>
<thead>
<tr>
<th>Subset of items in cluster</th>
<th>Cluster words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entrapment, Mission: Impossible III, Zombie, Snake Eyes, Starsky &amp; Hutch, New England Patriots vs. Minnesota Vikings, I Am Legend, Chaos</td>
<td>action, good, character, thought, story, plot, scene, expected, average, movies, game, scenes, lack, massive, destruction, entertained, suspenseful, audience, seats, batman</td>
</tr>
<tr>
<td>Gargantua, Random Hearts, Chocolate: Deep Dark Secrets, Blackout, The Ventures of Marguerite, Irresistible, Ghosts of Girlfriends Past, Youth Without Youth</td>
<td>like, good, bad, time, movies, people, acting, plot, watch, horror, watching, worst, scenes, pretty, awful, effects, scene, characters, thought, story, actors, worse, films, terrible, special, lot, fun</td>
</tr>
</tbody>
</table>

Table 1: Discovered clusters of items and associated topics for IMDb.

3. EXPERIMENTS

To extensively test our model, we select four datasets about movies, beer, businesses, and food. All four datasets come from different websites and communities, thus capturing different styles and patterns of online ratings and reviews. We evaluate performance of rating prediction based on RMSE, and review text prediction based on perplexity. An overview of our results can be seen in Figure 1, and detailed results for IMDb is shown in Table 1(d). Complete results for all four datasets are provided in the complete paper. We see PACO outperforms HFT and JMARS in rating prediction and achieves nearly as high quality review prediction as JMARS, while being far faster and simpler.

In addition to quantitatively evaluating our method, we also want to empirically demonstrate that the patterns surfaced would be useful to the human eye. We see PACO is able to find meaningful item clusters (Table 1), learn item-specific words (Table 4), and predict words matching the sentiment of the predicted rating (Table 3).}

4. CONCLUSION

We presented PACO, an additive novel Poisson co-clustering algorithm for explainable recommendations that is fast, succinct, interpretable and showed competitive results with state-of-the-art joint models.

Acknowledgements: This research was supported by funds from Google, a Facebook Fellowship, and the National Science Foundation under Grant No. DGE-1252522, CNS-1314632 and IIS-1408924. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, or other funding parties.

References