A detailed derivation of the context aware coupon recommendation

Part I, the context categorization problem

Our target is to optimize , which is a matrix, where are the number of contextual categories and features respectfully, that minimize the loss function as follows:



The first part of the loss function  minimizes the negative log-likelihood, the second part minimizes the KL-divergence to a prior knowledge, the third part  is a regularizer.

For each contextual base ,we keep for each part only terms related to , which leaves:

, for all reviews that are labeled as the k-th context

The probability of a review being labeled as the k-th context is: 

We have



Therefore



Similarly, we keep only terms related to in part 2, which leaves:



Since 



For 



Let





Part2 Preference Quantification Model

Given the observed commodity features in each review e, our target is to optimize , and the corresponding commodity feature distribution c, for each context x, by minimizing the loss function as follows:



Where  is predicted by the context categorization model of a review being positive labeled as context x.



