Correlation-Sensitive Next-Basket Recommendation

Tuo Liu

August 15, 2022

1 Main Contribution

1. The authors are the first to consider correlations among predicted items for this problem.
2. The authors further describe a novel hierarchical network architecture called Basket Sequence Correlation Network (codenamed Beacon), which learns the representations of each basket leading to the overall representation of a basket sequence that could be used for next-basket prediction.
3. The authors conduct extensive experiments on three real-life datasets of different domains. The results show that Beacon's modeling of item correlations produce significant improvements over baselines.

2 Model

The framework Beacon consists of three main components:

1. correlation-sensitive basket encoder: Taking a basket sequence and correlation matrix as input, the basket encoder captures intra-basket item correlations and produce correlation-sensitive basket representations.
2. basket sequence encoder: The sequence of basket representations is further fed into a sequence encoder to capture inter-basket sequential associations.
3. correlation-sensitive predictor: The output from the sequence encoder, together with the correlation matrix, are employed by the predictor to produce the correlation-sensitive next basket.

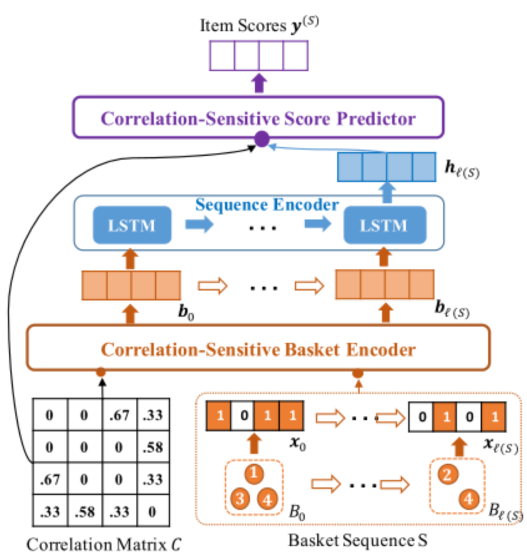

Figure 1: Structure

2.1 Integrating Temporal Dynamics

The encoder is as follow:

$$\mathbf{z}_t = \mathbf{x}_t \circ \boldsymbol{\omega} + \operatorname{ReLU}\left(\mathbf{x}_t C - \eta \mathbf{1}\right)$$

where x_t is the basket vector (element be 0 or 1), C is the input correlation matrix, η is a learnable scalar parameter to filter out weak correlations. Subsequently, z_t is fed into a fully-connected layer to infer a latent L-dimension basket representation b_t, as follows:

$$\mathbf{b}_t = \operatorname{ReLU}\left(\mathbf{z}_t \Phi + \phi\right)$$

where Φ and ϕ stand for weight and bias respectively.

2.2 Basket Sequence Encoder

The sequence encoder employs a RNN to capture the sequential associations in basket sequences. The hidden output is as follow:

$$\mathbf{h}_t = \tanh\left(\mathbf{b}_t \Psi + \mathbf{h}_{t-1} \Psi' + \boldsymbol{\psi}\right)$$

while Beacon adopts LSTM units, it is flexible to plug in other recurrent units.

2.3 Correlation-Sensitive Score Predictor

The predictor aims to derive a score for each item based on both the inter-basket sequential associations and intra-basket correlation associations. Let $\mathbf{h}_{\ell(S)}$ be the last hidden output of sequence S via the sequence encoder. Thus, the sequential signal for item recommendation given sequence S can be estimated by the following:

$$\mathbf{s}^{(S)} = \sigma\left(\mathbf{h}_{\ell(S)} \Gamma\right)$$

In order to recommend a basket with correlated items, we further aggregate the sequential signal with item importance and correlative associations. a straightforward solution is $\left(\mathbf{s}^{(S)} \circ \boldsymbol{\omega} + \mathbf{s}^{(S)} C\right)$. However, in this formulation, the intra-basket correlative association often dominates and masks the inter-basket sequential associations. Thus, the authors adopt the following predictor, such that the trade-off between correlative and sequential associations can be tuned:

$$\mathbf{y}^{(S)} = \alpha\left(\mathbf{s}^{(S)} \circ \boldsymbol{\omega} + \mathbf{s}^{(S)} C\right) + (1-\alpha)\mathbf{s}^{(S)}$$

2.4 Next-Basket Recommendation

Since the size of the next basket is unknown and is often noncritical in a recommendation setting, the authors form the next basket by taking K items with the highest scores in $y(S)$.

2.5 Learning Strategy

For each training sequence S, we remove its last basket. The goal is to make sure that the predicted scores should align well with the ground truth next basket. To this end, we favor the adopted items in the ground truth basket, and at the same time penalize other negative items, as follow:

$$\begin{aligned}\mathcal{L}(S) = &- \frac{1}{\left|B_{\ell(S)}\right|} \sum_{i \in B_{\ell(S)}} \log \sigma\left(\mathbf{y}_i^{(S')}\right) \\ &- \frac{1}{\left|V \backslash B_{\ell(S)}\right|} \sum_{j \in V \backslash B_{\ell(S)}} \log\left(1 - \sigma\left(\mathbf{y}_j^{(S')} - \mathbf{y}_m^{(S')}\right)\right)\end{aligned}$$

3 Experiments

The paper experiments with two benchmark datasets: **TaFeng, Delicious, Foursquare**. The baselines include **POP, MC, MCN, DREAM, BSEQ, Triple2vec**. The results are as follows:

Dataset	Model	L	H	F1@K (%) @5	F1@K (%) @10	HLU
TaFeng	*POP*	-	-	4.66	4.02	6.64
	MC	-	-	4.11	3.61	5.78
	MCN	8	-	4.56	4.02	6.34
	DREAM	8	-	5.85	4.90	6.96
	BSEQ	32	16	4.48	4.04	6.34
	triple2vec	64	-	4.66	3.88	4.85
	Beacon	8	64	**6.36**†	**5.26**†	**7.83**†
Delicious	*POP*	-	-	3.88	4.04	6.05
	MC	-	-	4.27	4.59	6.52
	MCN	32	-	4.20	4.59	6.50
	DREAM	32	-	3.13	3.47	4.93
	BSEQ	64	32	3.86	3.97	5.95
	triple2vec	32	-	3.76	4.04	5.16
	Beacon	64	64	**4.93**†	**5.47**†	**7.76**†
Foursquare	*POP*	-	-	2.73	2.90	4.84
	MC	-	-	3.58	3.43	5.53
	MCN	64	-	3.09	2.89	5.08
	DREAM	64	-	2.84	3.00	4.98
	BSEQ	64	32	2.80	2.89	4.82
	triple2vec	64	-	2.73	2.90	4.53
	Beacon	64	64	**3.61**	**3.59**†	**6.32**†

Figure 2: Performance

1. For TaFeng, popularity seems to be an important factor since POP performs better than MC, MCN, BSEQ and triple2vec. Beyond popularity, DREAM and Beacon show advantages in capturing associations between basket items. Yet, Beacon is the best-performing model.

2. For Delicious, Markov-based models (MC and MCN) do better than other baselines. It might imply that items in a testing basket are strongly dependent on the most recent basket. The modeling of basket-oriented associations in DREAM and triple2vec is not helpful to improve the performance. In contrast, Beacon shows a significant improvement over these models across the three measurements, which we attribute to the advantage of modeling correlations effectively.

3. For Foursquare, we witness a similar observation as Delicious, where Beacon outperforms the baselines significantly.

4 My Idea

In this paper, the authors propose a model named Beacon that utilizes the correlation information to enhance the representation of individual baskets as well as the overall basket sequence. It is really novel because in the past people always only use the history information and ignore the correlation information within the basket. With the correlation information, the performance of the model indeed becomes better.