A WORKER FEEDBACK

Figure 9 visualizes worker feedback towards our task, where (a) shows the feedback regarding the difficulty on identifying bundles and naming the corresponding intents for the two batches; and (b-c) depict the general feedback for the two batches.

B WORKER BASIC INFORMATION

Figure 10 shows the distribution of workers' age, education, country, occupation, gender and shopping frequency for the two batches.

C BUNDLE DATA DISTRIBUTION

Figure 11(a) shows the item distribution regarding the number of affiliated bundles. Figure 11(b) shows the user distribution regarding the number of interacted bundles.

D THE STRATEGY FOR PRE-TRAINING

As observed in Table 2, the user-bundle interaction matrix is quite sparse in the three domains. To ensure the quality of learned representations, we adopt BPRMF to pre-train item representations to initialize all methods; meanwhile the bundle representation is initialized with the averaged representation of items inside this bundle. To avoid data leakage, we sample extra user-item interactions from each domain with the same time span as bundle collection (i.e. July 2013 - July 2014) for pre-training. In particular, for each item in Table 2, we first find out all users who have interacted with it excluding those in Table 2, and then leverage these users' records to construct the user-item interaction matrix for pre-training. By doing so, the constructed interaction data covers all items but excludes all users in Table 2, therefore, do not leak any interactions in our collected data in Table 2. The statistics of datasets for pre-training across the three domains are summarized in Table 6.

Table 6: Statistics of datasets for pre-training.

	#Users	#Items	#Interactions
Electronic	745,911	178,443	1,951,397
Clothing	223,571	236,387	749,423
Food	143,390	54,323	325,169

E PARAMETER TUNING AND SETTINGS FOR BUNDLE DETECTION

A grid search in $\{0.0001, 0.001, 0.01\}$ is applied to find out the optimal settings for $support$ and $confidence$, and both are set as 0.001 across the three domains.

F PARAMETER TUNING AND SETTINGS FOR BUNDLE COMPLETION

The dimension (d) of item and bundle representations for all methods is 20. Grid search is adopted to find out the best settings for other key parameters. In particular, learning rate (η) and regularization coefficient (λ) are searched in $\{0.0001, 0.001, 0.01\}$; the number of neighbors (K) in ItemKNN is searched in $\{10, 20, 30, 50\}$; the weight of KL divergence (α) in VAE is searched in $\{0.001, 0.01, 0.1\}$; and the batch size is searched in $\{64, 128, 256\}$. The optimal parameter settings are shown in Table 7.

Table 7: Parameter settings for bundle completion ($d = 20$).

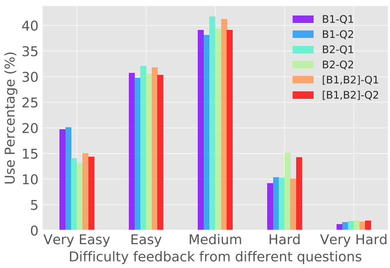
	Electronic	Clothing	Food
ItemKNN	$K = 10$	$K = 10$	$K = 10$
BPRMF	$\eta = 0.0001$ $\lambda = 0.001$ $neg\_sample = 2$ $batch\_size = 128$	$\eta = 0.0001$ $\lambda = 0.01$ $neg\_sample = 2$ $batch\_size = 128$	$\eta = 0.01$ $\lambda = 0.01$ $neg\_sample = 2$ $batch\_size = 128$
mean-VAE	$\eta = 0.0001$ $\lambda = 0.001$ $\alpha = 0.01$ $hid\_layers = [100, 50]$ $dropout = 0.5$ $batch\_size = 64$	$\eta = 0.0001$ $\lambda = 0.0001$ $\alpha = 0.001$ $hid\_layers = [100, 50]$ $dropout = 0.5$ $batch\_size = 128$	$\eta = 0.0001$ $\lambda = 0.0001$ $\alpha = 0.001$ $hid\_layers = [100, 50]$ $dropout = 0.5$ $batch\_size = 64$
concat-VAE	$\eta = 0.0001$ $\lambda = 0.01$ $\alpha = 0.001$ $dropout = 0.5$ $batch\_size = 128$	$\eta = 0.001$ $\lambda = 0.001$ $\alpha = 0.1$ $dropout = 0.5$ $batch\_size = 64$	$\eta = 0.0001$ $\lambda = 0.0001$ $\alpha = 0.001$ $dropout = 0.5$ $batch\_size = 64$

Table 8: Parameter settings for bundle ranking ($d = 20$).

	Electronic	Clothing	Food
ItemKNN	$K = 10$	$K = 10$	$K = 10$
BPRMF	$\eta = 0.0001$ $\lambda = 0.001$ $neg\_sample = 2$ $batch\_size = 128$	$\eta = 0.0001$ $\lambda = 0.01$ $neg\_sample = 2$ $batch\_size = 128$	$\eta = 0.0001$ $\lambda = 0.0001$ $neg\_sample = 2$ $batch\_size = 128$
DAM	$\eta = 0.01$ $neg\_sample = 1$ $dropout = 0.5$	$\eta = 0.01$ $neg\_sample = 1$ $dropout = 0.5$	$\eta = 0.01$ $neg\_sample = 1$ $dropout = 0.5$
AttList	$\eta = 0.001$ $neg\_sample = 2$ $\#bundles/user = 5$ $\#items/bundle = 10$ $D = 100$ $dropout = 0.5$ $batch\_size = 64$	$\eta = 0.0001$ $neg\_sample = 2$ $\#bundles/user = 5$ $\#items/bundle = 10$ $D = 50$ $dropout = 0.5$ $batch\_size = 128$	$\eta = 0.001$ $neg\_sample = 2$ $\#bundles/user = 5$ $\#items/bundle = 10$ $D = 50$ $dropout = 0.5$ $batch\_size = 256$
GCN	$\eta = 0.01$ $\lambda = 0.01$ $neg\_sample = 1$ $msg\_dropout = 0.3$ $node\_dropout = 0$ $prop\_layers = 2$ $batch\_size = 2048$	$\eta = 0.001$ $\lambda = 0.0001$ $neg\_sample = 1$ $msg\_dropout = 0.5$ $node\_dropout = 0$ $prop\_layers = 2$ $batch\_size = 2048$	$\eta = 0.01$ $\lambda = 0.0001$ $neg\_sample = 1$ $msg\_dropout = 0.5$ $node\_dropout = 0$ $prop\_layers = 2$ $batch\_size = 2048$
BGCN	$\eta = 0.001$ $\lambda = 0.001$ $neg\_sample = 1$ $msg\_dropout = 0.1$ $node\_dropout = 0$ $prop\_layers = 2$ $batch\_size = 2048$	$\eta = 0.01$ $\lambda = 0.01$ $neg\_sample = 1$ $msg\_dropout = 0$ $node\_dropout = 0$ $prop\_layers = 2$ $batch\_size = 2048$	$\eta = 0.01$ $\lambda = 0.001$ $neg\_sample = 1$ $msg\_dropout = 0.1$ $node\_dropout = 0.1$ $prop\_layers = 2$ $batch\_size = 2048$

G PARAMETER TUNING AND SETTINGS FOR BUNDLE RANKING

The dimension (d) of representations is set as 20. We apply a same grid search for η, λ, K and batch size as in bundle completion. Besides, the predictive layer D for AttList is searched from $\{20, 50, 100\}$; the node and message dropout rate for GCN and BGCN is searched in $\{0, 0.1, 0.3, 0.5\}$. As the training complexity for GCN and BGCN is quite high, we set the batch size as 2048 as suggested by the original paper. The optimal parameter settings are presented in Table 8. Note that the parameter settings for BGCN is the version without pre-training (i.e. $\text{BGCN}_{w/o\ pre}$).

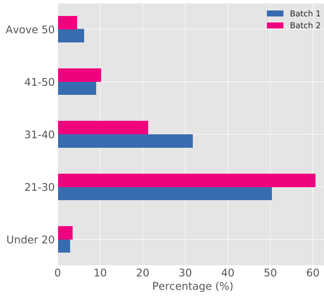


(a) Difficulty feedback for two batches

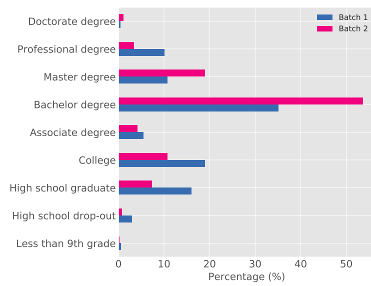
(b) General feedback for the first batch

(c) General feedback for the second batch

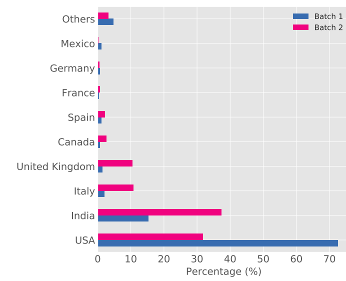
Figure 9: Feedback from workers for the two batches.



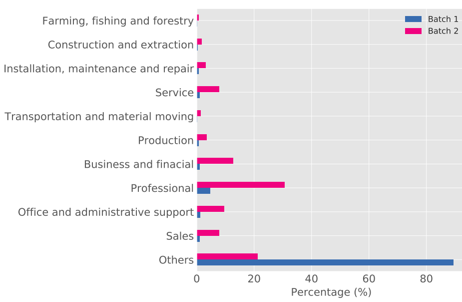
(a) Distribution of age



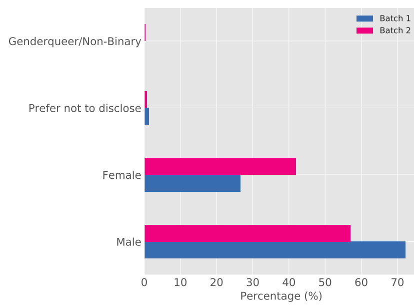
(b) Distribution of education



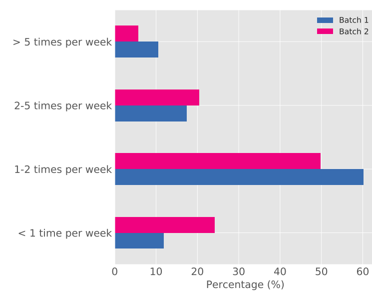
(c) Distribution of country



(d) Distribution of occupation

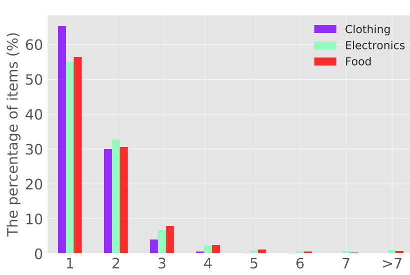


(e) Distribution of gender

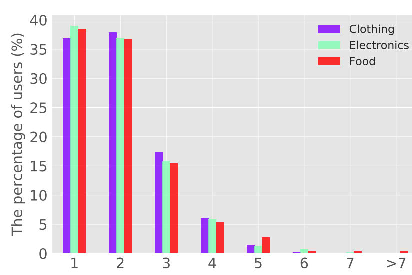


(f) Distribution of shopping frequency

Figure 10: Worker basic information in the first and second batch. In particular, 'Others' in the country distribution includes Argentina, Australia, Anguilla, Netherlands, Albania, Georgia, Tunisia, Belgium, Armenia, Guinea, Austria, Switzerland, Iceland, Lithuania, Egypt, Venezuela, Bangladesh, American Samoa, Vanuatu, Colombia, United Arab Emirates, Ashmore and Cartier Island, Estados Unidos, Wales, Turkey, Angola, Scotland, Philippines, Iran and Bahamas.



(a) Item-bundle distribution



(b) User-bundle distribution

Figure 11: Item-bundle and user-bundle distribution across the three domains.