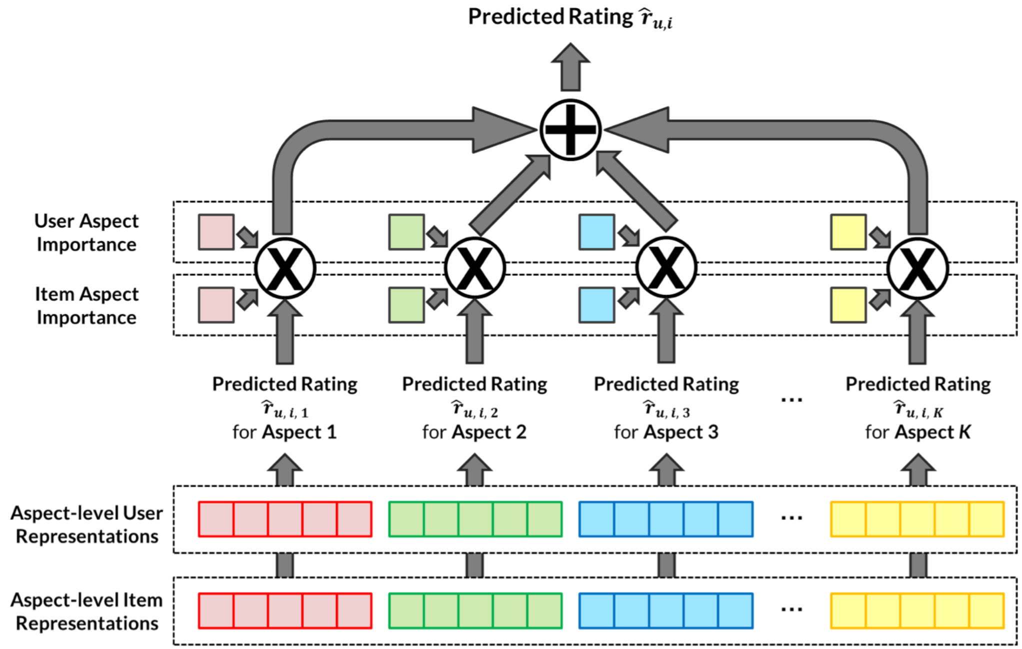


ANR: Aspect-based Neural Recommender

Jin Yao Chin, **Kaiqi Zhao**, **Shafiq Joty**, and **Gao Cong**

School of Computer Science and Engineering
Nanyang Technological University, Singapore

NANYANG TECHNOLOGICAL UNIVERSITY
SINGAPORE

Outline

- ▷ Problem Formulation & Existing Work
- ▷ Proposed Model: Aspect-based Neural Recommender
- ▷ Experimental Results
- ▷ Future Work & Conclusion

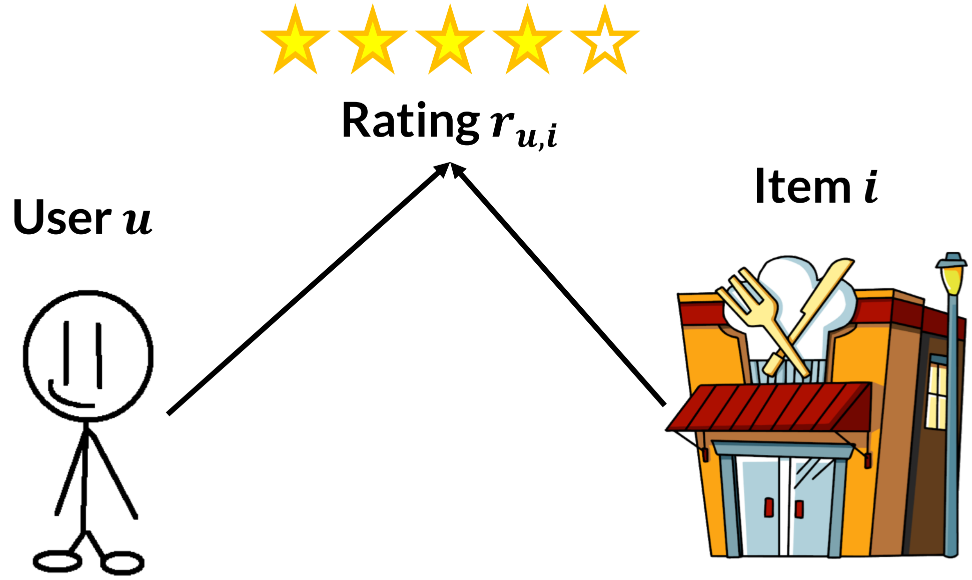
1.

Overview

General Recommendation

For each user u, we would like to estimate the rating $\hat{r}_{u,i}$ for any **new item** i

- **Explicit Feedback Matrix** $\boldsymbol{R} \in \mathbb{R}^{N \times M}$
  - N users, M items
  - $r_{u,i} = \{1, \dots, 5\}$ if user u has interacted with item i, 0 otherwise



- Recommend **new items** that the user would rate highly

Recommendation with Reviews

Solid MI

By Zimmer on September 2, 2018

Format: Blu-ray

No doubt one of if not the best movie released this year and, just my IMO, in the top 3 Mission films. However im not sure it is quite deserving of the high RT rating it received. It does drag a bit in the second act when Solomon Lane is introduced again. The film needed a truly great scene stealing villain IMO to compete with the great action and Cruise's stunts, and Lane just isnt that interesting. Much has been said of Cavill and his amazing moustache and he's decent but a bit wooden. Great physical presence though. Cruise is solid as usual. Really really enjoyed the first act and the action scenes toward the end were great. The score by Lorne Balfe might just be the best MI score yet. Should have cut the running time a bit tho

- **Assumption**: Each user-item interaction contains a **textual review**
  - Readily available in many e-commerce and review websites (E.g. *Yelp*, *Amazon*, etc)

- A **complete user-item interaction**: $(u,\ i,\ r_{u,i},\ d_{u,i})$

Rating

Review

“Problems” with Reviews

Solid MI

By Zimmer on September 2, 2018

Format: Blu-ray

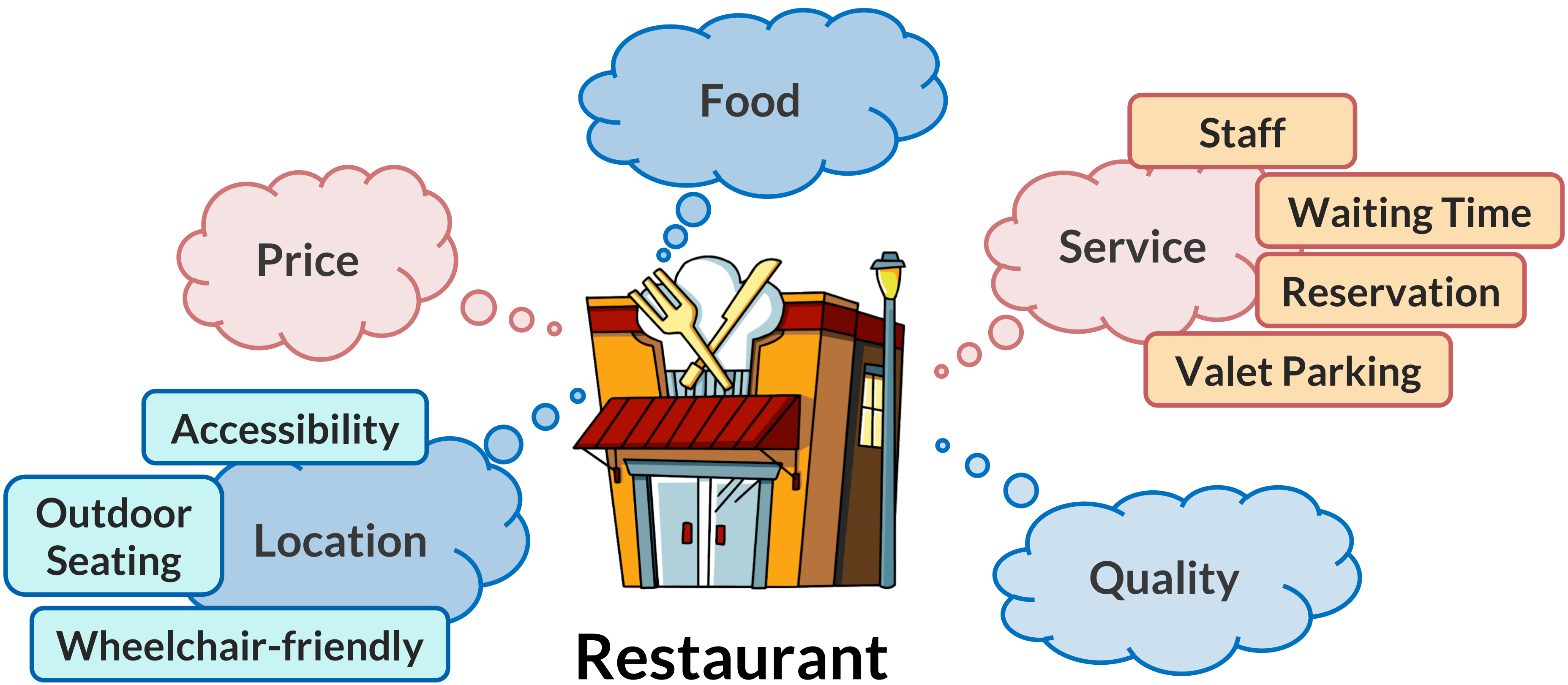
No doubt one of if not the best movie released this year and, just my IMO, in the top 3 Mission films. However im not sure it is quite deserving of the high RT rating it received. It does drag a bit in the second act when Solomon Lane is introduced again. The film needed a truly great scene stealing villain IMO to compete with the great action and Cruise’s stunts, and Lane just isnt that interesting. Much has been said of Cavill and his amazing moustache and he’s decent but a bit wooden. Great physical presence though. Cruise is solid as usual. Really really enjoyed the first act and the action scenes toward the end were great. The score by Lorne Balfe might just be the best MI score yet. Should have cut the running time a bit tho

1. **Not all parts of the review are equally important!**
   - E.g. “The restaurant is located beside a old-looking post office” may not be correlated with the overall user satisfaction

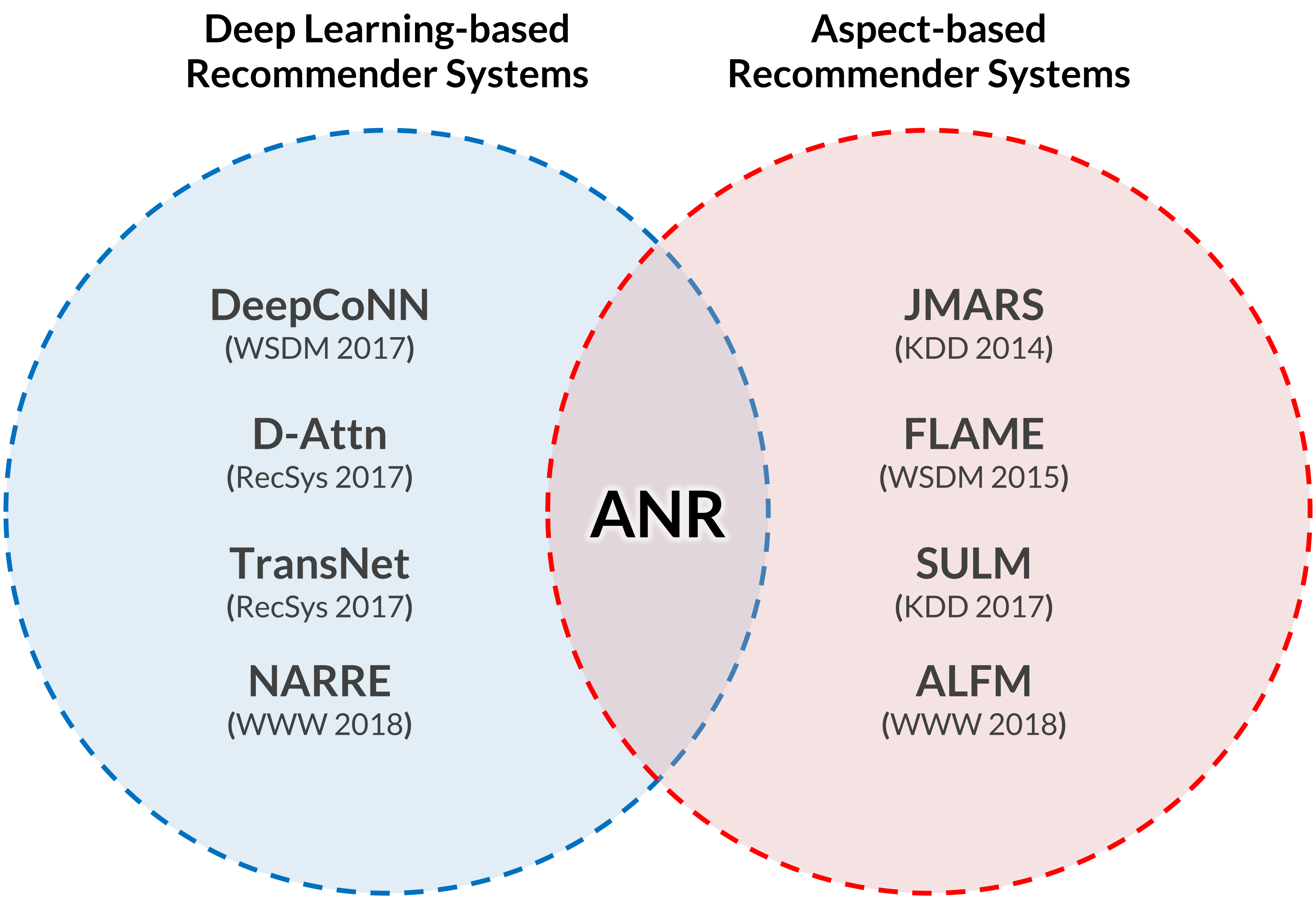
2. **Each review may cover multiple “aspects”**
   - **Review Length**: Around 100 to 150 words in general
   - Users may describe about various item properties

What is an Aspect?

- A ***high-level semantic concept***
- Encompasses a specific facet of *item properties* for a *given domain*



Existing Work & Our Model



Existing Work & Our Model

Deep Learning-based Recommender Systems

- ✓ Capitalizes on the <u>strong representation learning capabilities</u> of neural networks
- × Less interpretable and informative

Aspect-based Recommender Systems

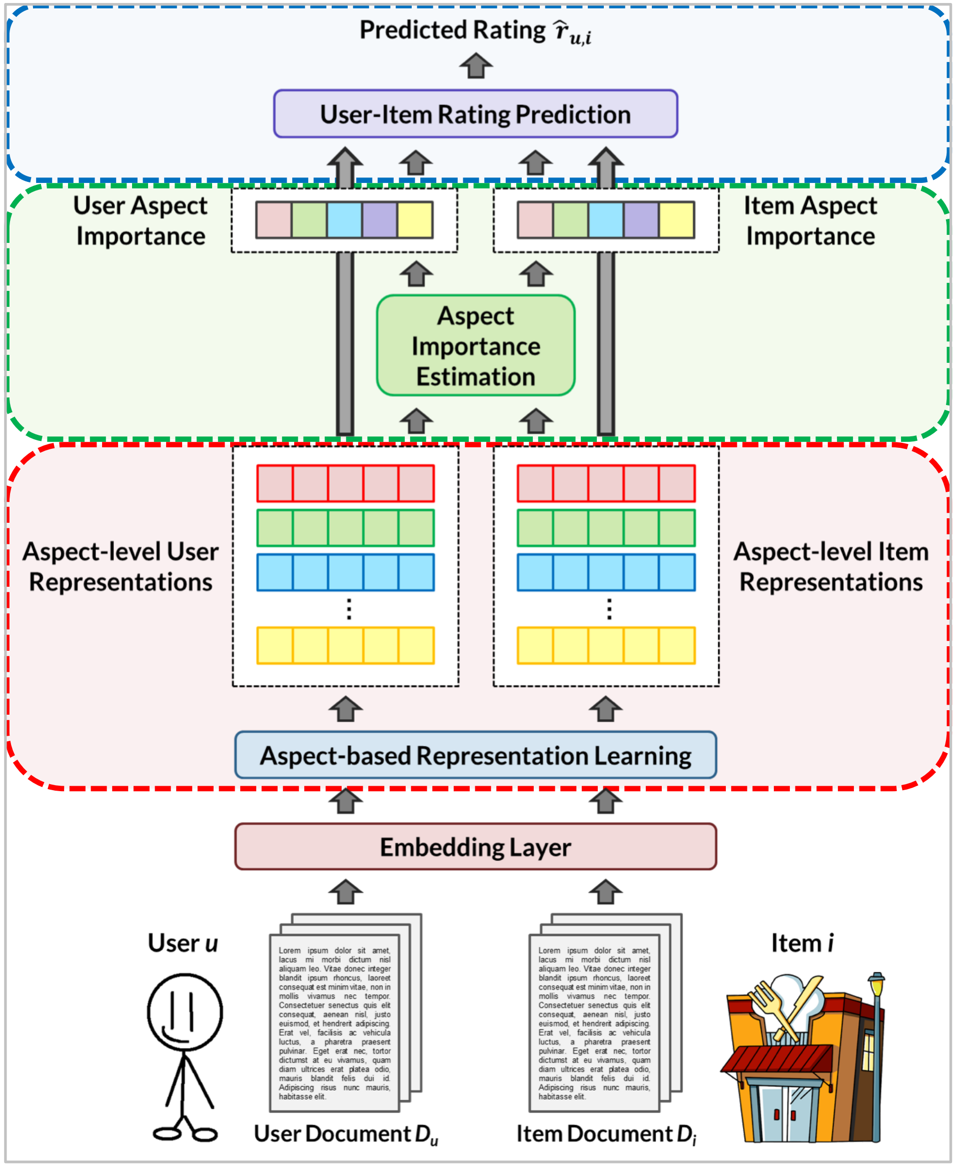
- ✓ More <u>interpretable & explainable recommendations</u>
- × May rely on existing Sentiment Analysis (**SA**) tools for the extraction of aspects and/or sentiments
- × Not self-contained
- × Performance can be limited by the quality of these SA tools

Our Model: Combines the strengths of these two categories of recommender systems

2.

Proposed Model

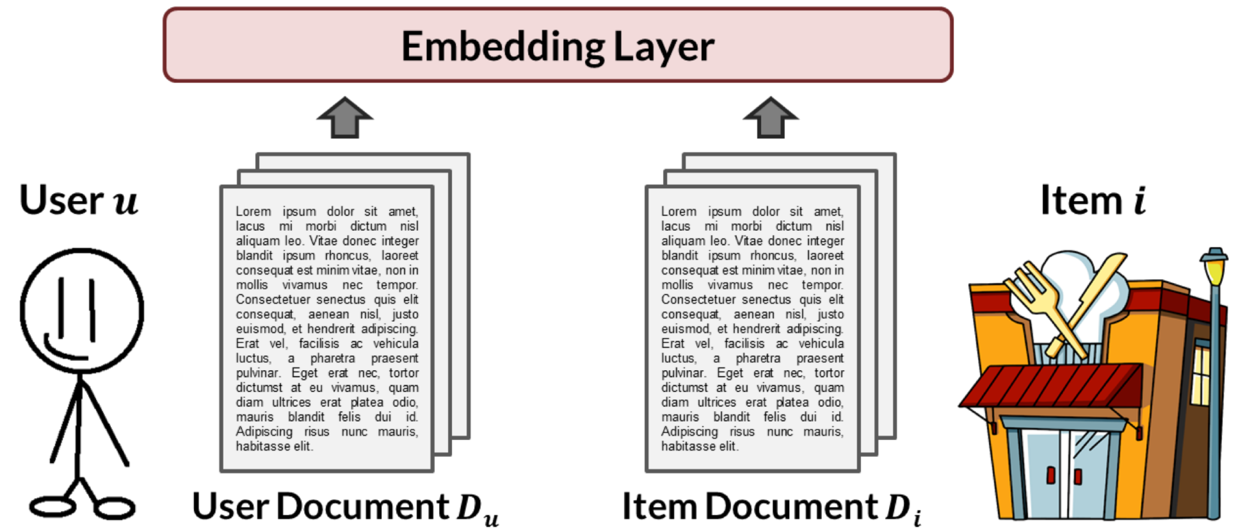
Our Proposed Model - ANR



Key Components

- ***Aspect-based Representation Learning*** to derive the aspect-level user and item latent representations
- Interaction-specific ***Aspect Importance Estimation*** for both the user and item
- ***User-Item Rating Prediction*** by effectively combining the aspect-level representations and importance

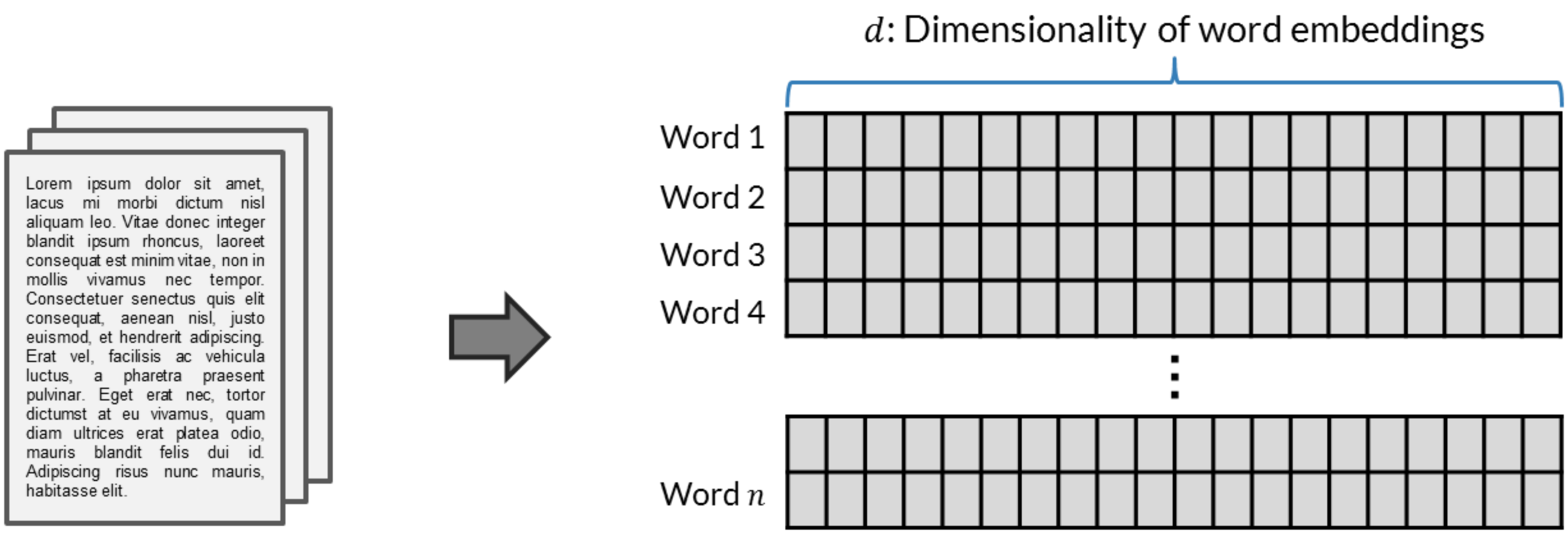
Input & Embedding Layer



Input

- ▷ Similar to existing deep learning-based methods
- ▷ **User document** D_u consists of the set of review(s) written by user u
- ▷ **Item document** D_i consists of the set of review(s) written for item i

Input & Embedding Layer



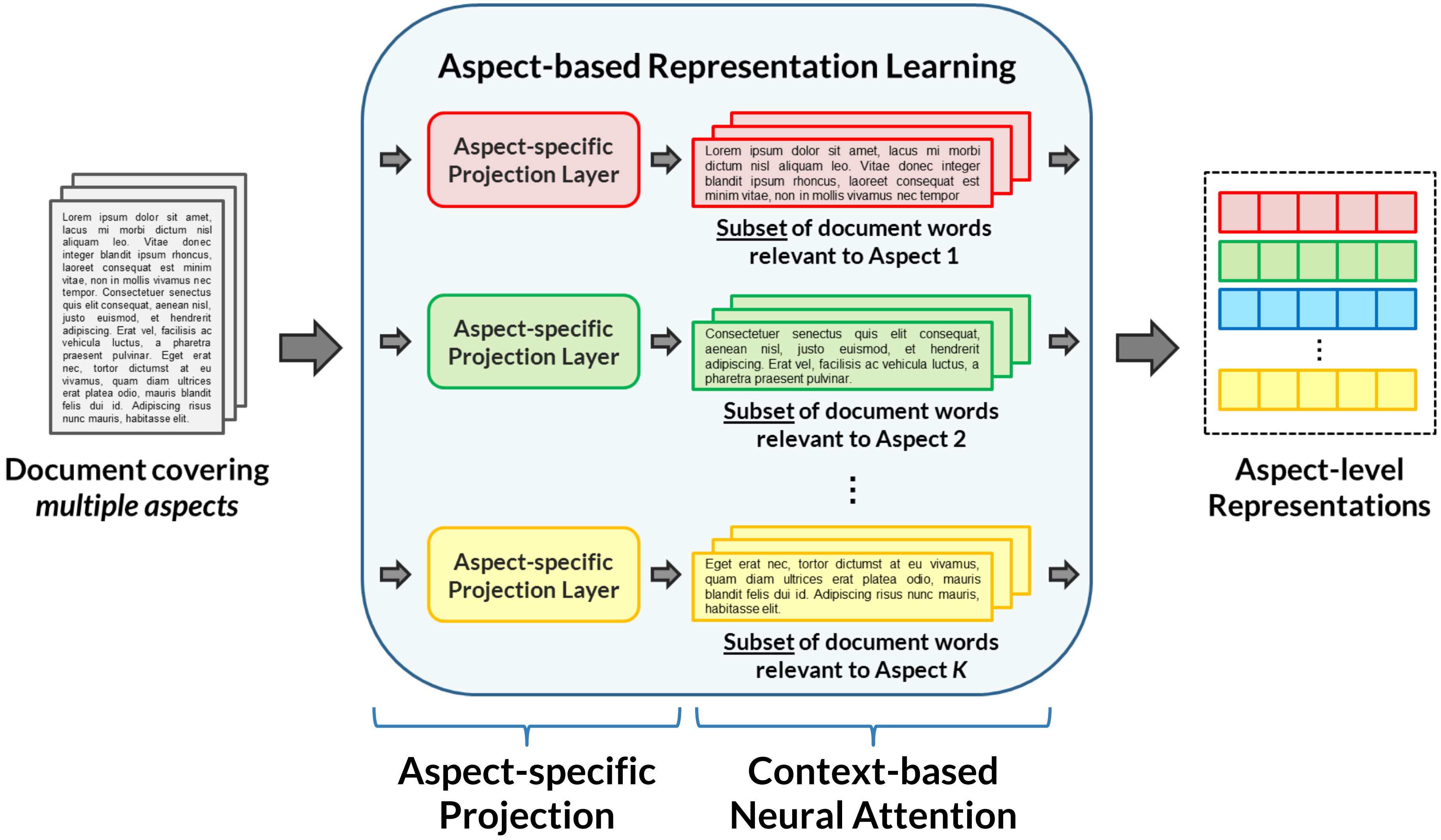
Document with
n words

Matrix with
n rows and d columns

Embedding Layer

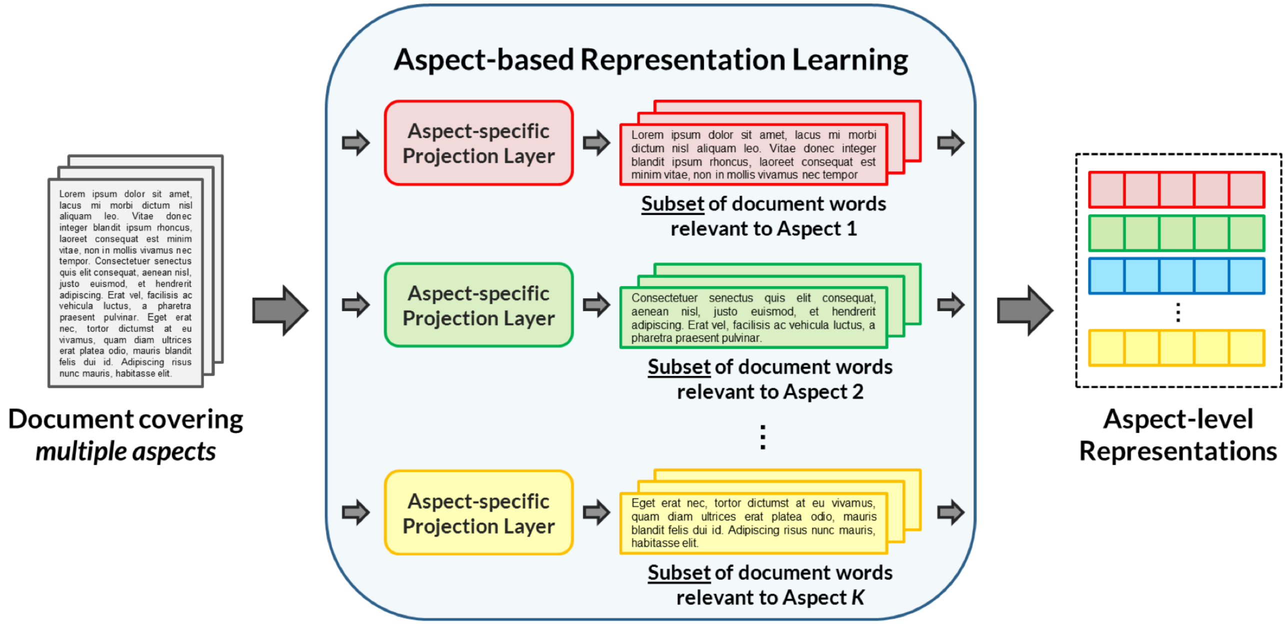
- ▷ Look-up operation in a embedding matrix (shared between users & items)
- ▷ ***Order*** and ***context*** of words within each document is preserved

Aspect-based Representations



Assumption: K aspects (Pre-defined Hyperparameter)

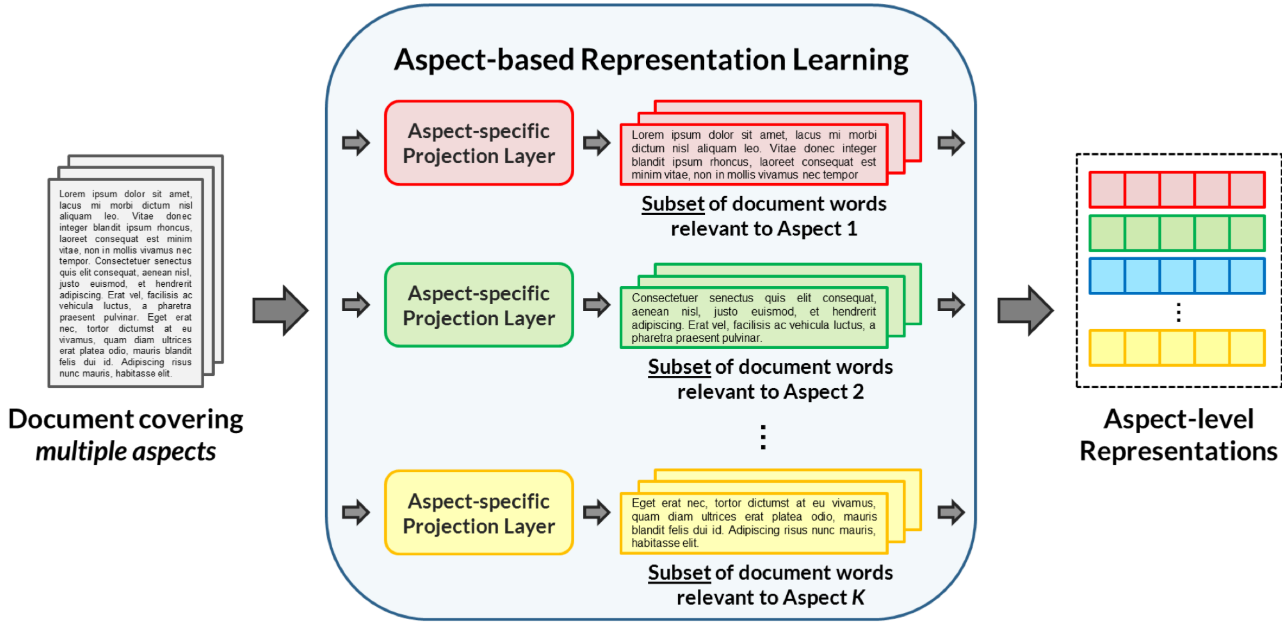
Aspect-based Representations



Aspect-specific Projections

- Semantic polarity of a word may vary for different aspects
- "The phone has a ***high*** storage capacity" ✓ ☺
- "The phone has extremely ***high*** power consumption" × ☹

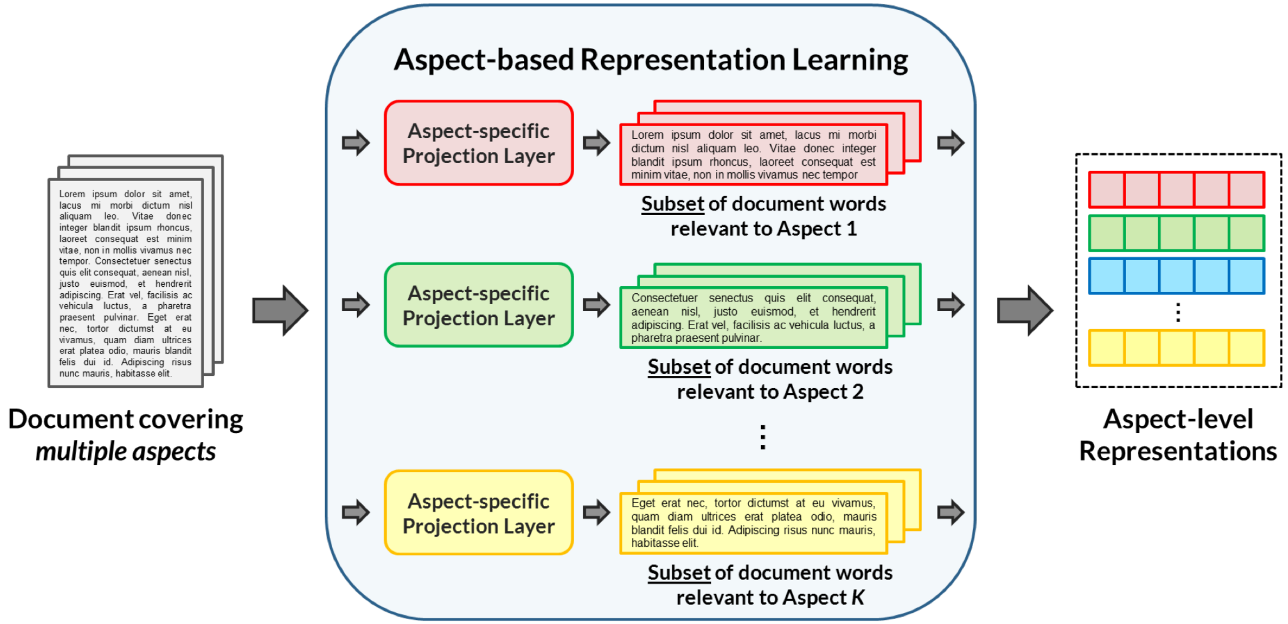
Aspect-based Representations



Context-based Neural Attention

- **Local Context**: Target word & its surrounding words
- **Word Importance**: Inner product of the word embeddings (within local context window) and the corresponding aspect embedding

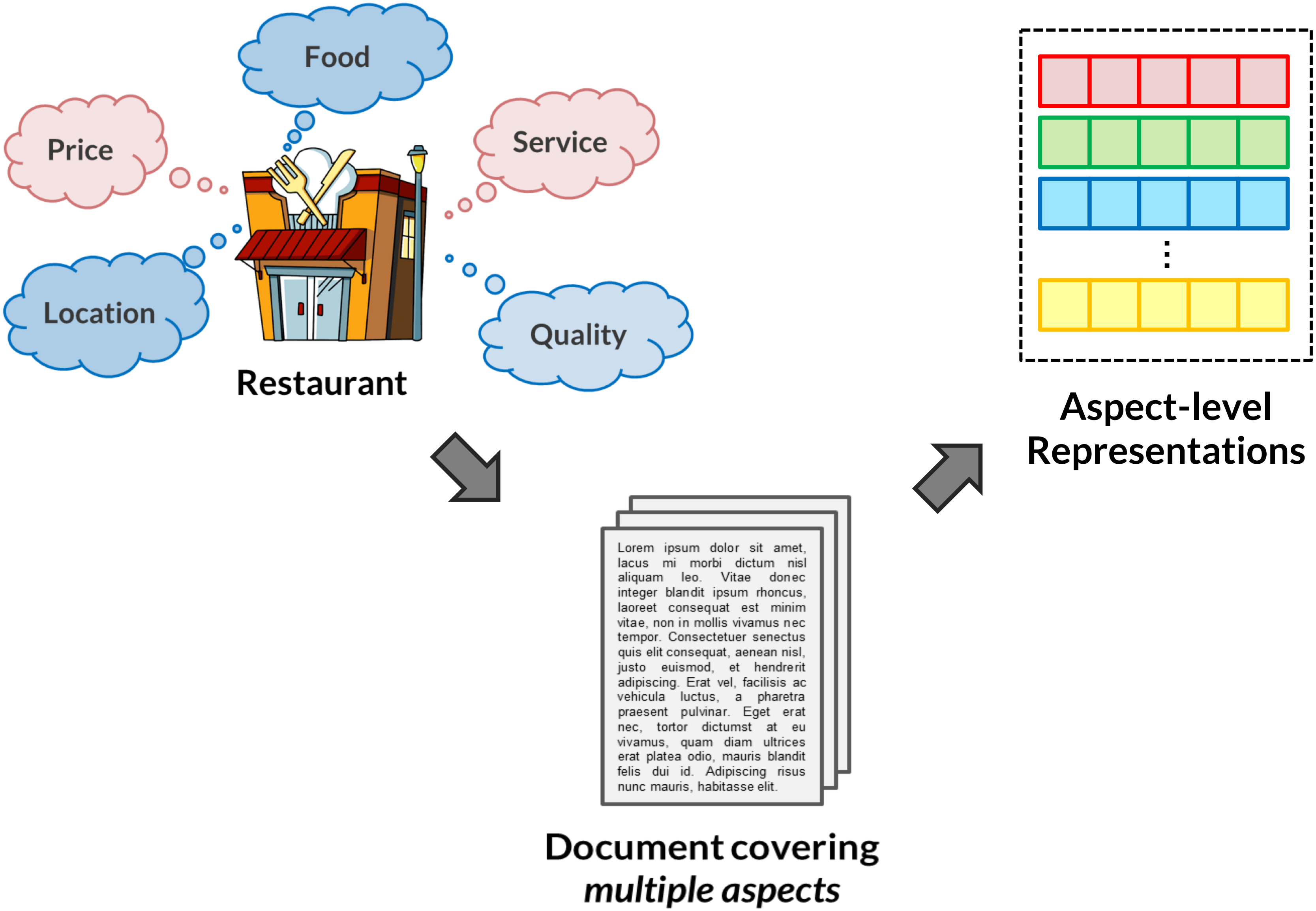
Aspect-based Representations



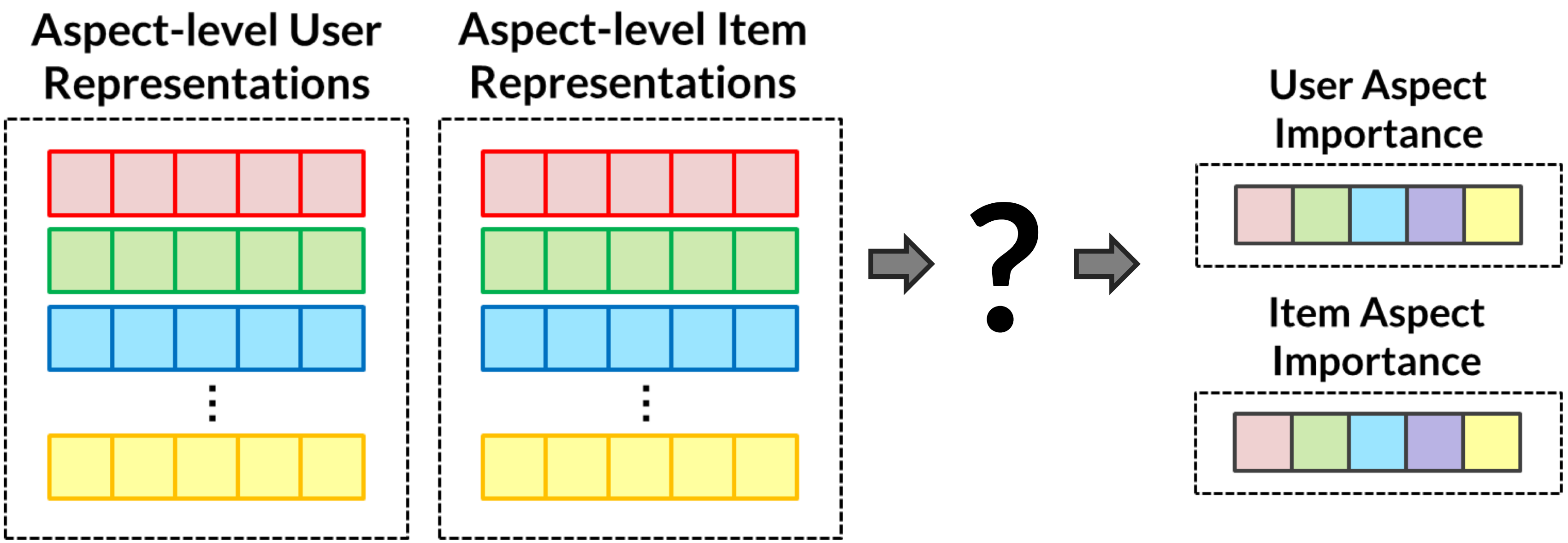
Aspect-level Representations

- **Weighted sum** of document words based on the *learned aspect-level word importance*
- Captures the same document from **multiple perspectives** by **attending** to different subsets of document words

Aspect-based Representations



User & Item Aspect Importance

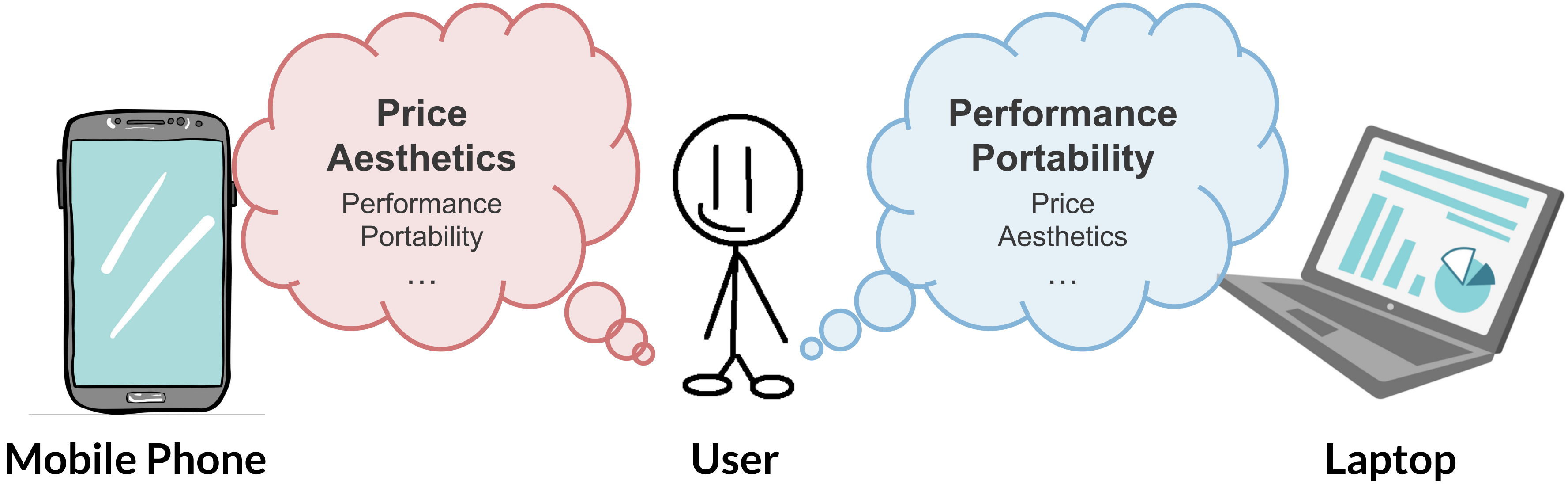


Goal: Estimate the user & item aspect importance for each user-item pair

- ▷ Based on **3 key observations**
- ▷ Extends the idea of **Neural Co-Attention** (i.e. Pairwise Attention)

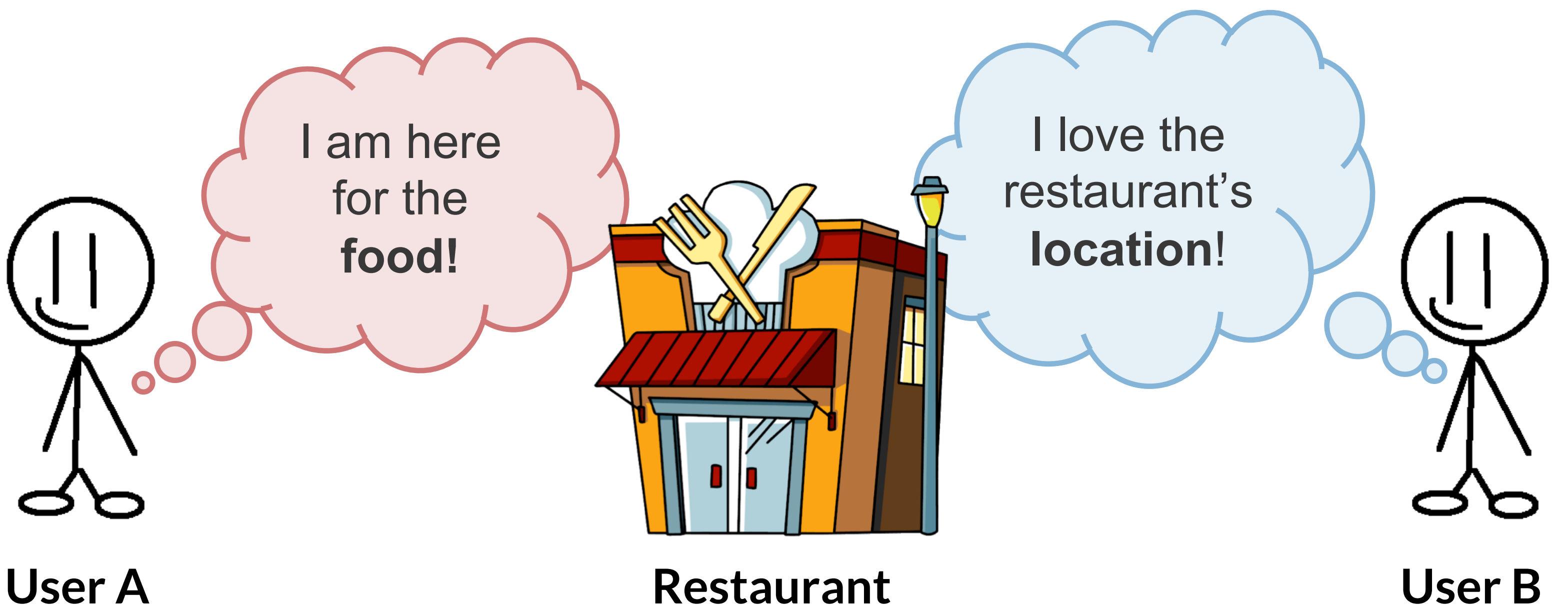
Dynamic Aspect-level Importance

1. A user's aspect-level preferences may change with respect to the target item
2. The same item may appeal differently to two different users
3. These aspects are often not evaluated separately/independently



Dynamic Aspect-level Importance

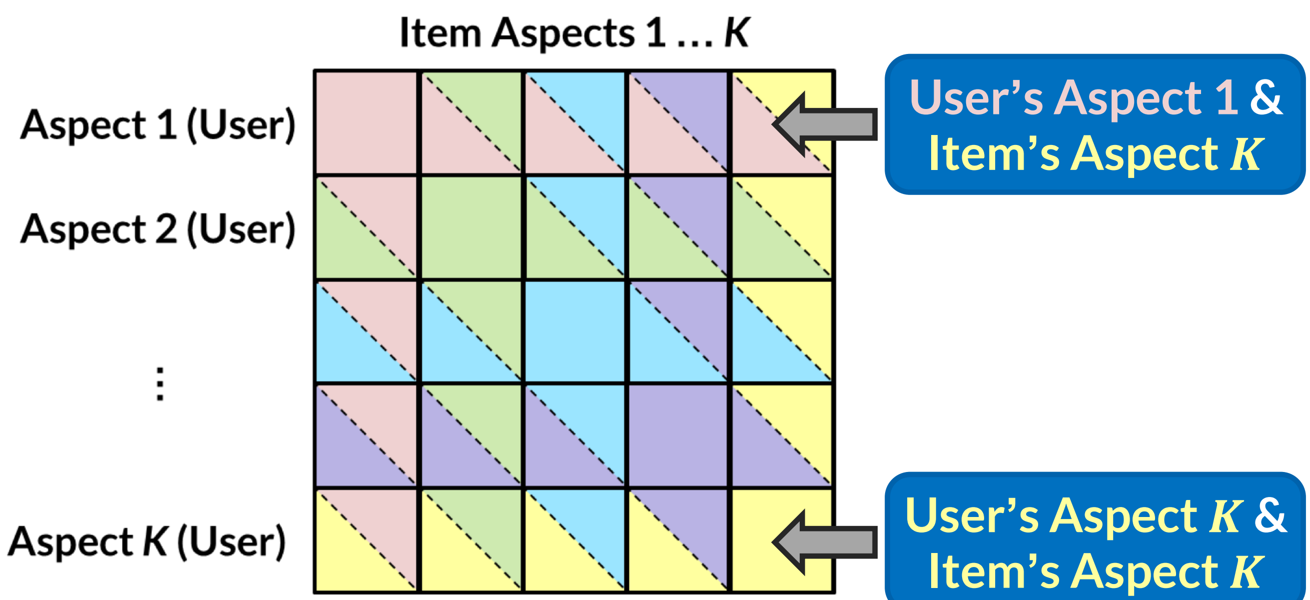
1. A user's aspect-level preferences may change with respect to the target item
2. **The same item may appeal differently to two different users**
3. These aspects are often not evaluated separately/independently



Dynamic Aspect-level Importance

1. A user's aspect-level preferences may change with respect to the target item
2. The same item may appeal differently to two different users
3. **These aspects are often not evaluated separately/independently**

Dynamic Aspect-level Importance



Affinity Matrix

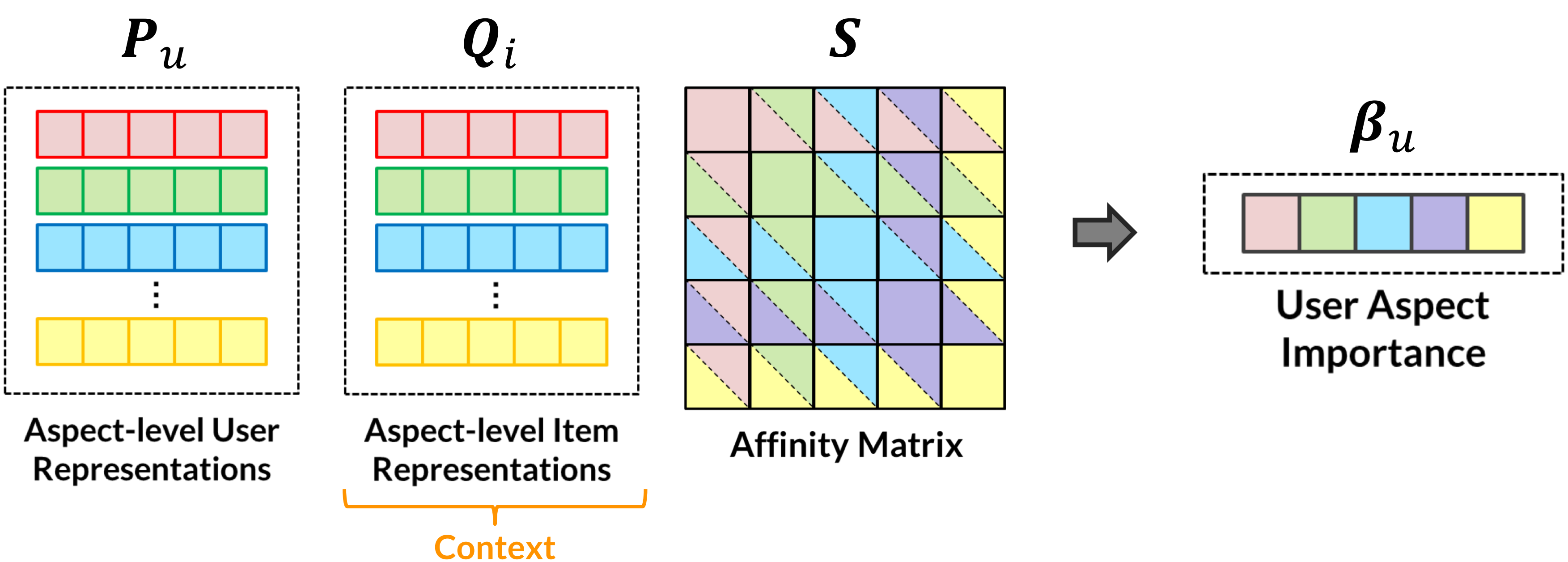
- Captures the **'shared similarity'** between the aspect-level representations
- Used as a **feature** for deriving the user & item aspect importance

Dynamic Aspect-level Importance

User Aspect Importance:

$$\boldsymbol{H}_u = \emptyset\left(\boxed{\boldsymbol{P}_u}\boldsymbol{W}_x + \boxed{\boldsymbol{S}^{\top}}\left(\boxed{\boldsymbol{Q}_i}\boldsymbol{W}_y\right)\right)$$

$$\boldsymbol{\beta}_u = softmax(\boldsymbol{H}_u\,\boldsymbol{v}_x)$$

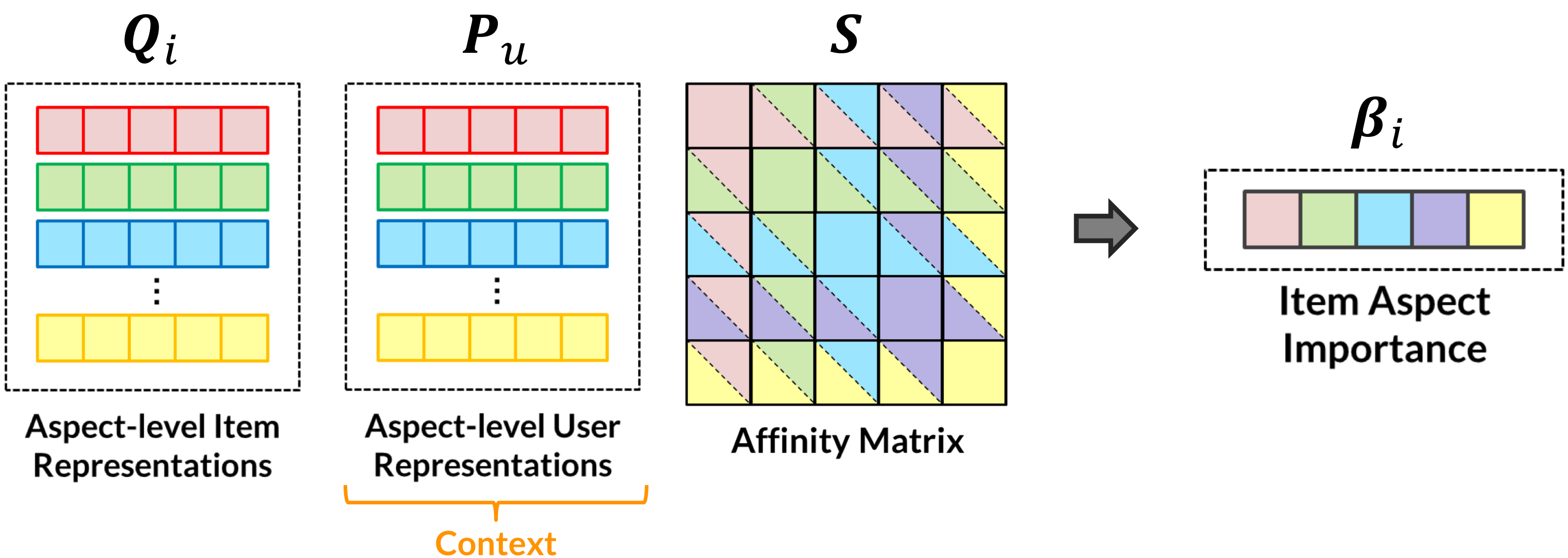


Dynamic Aspect-level Importance

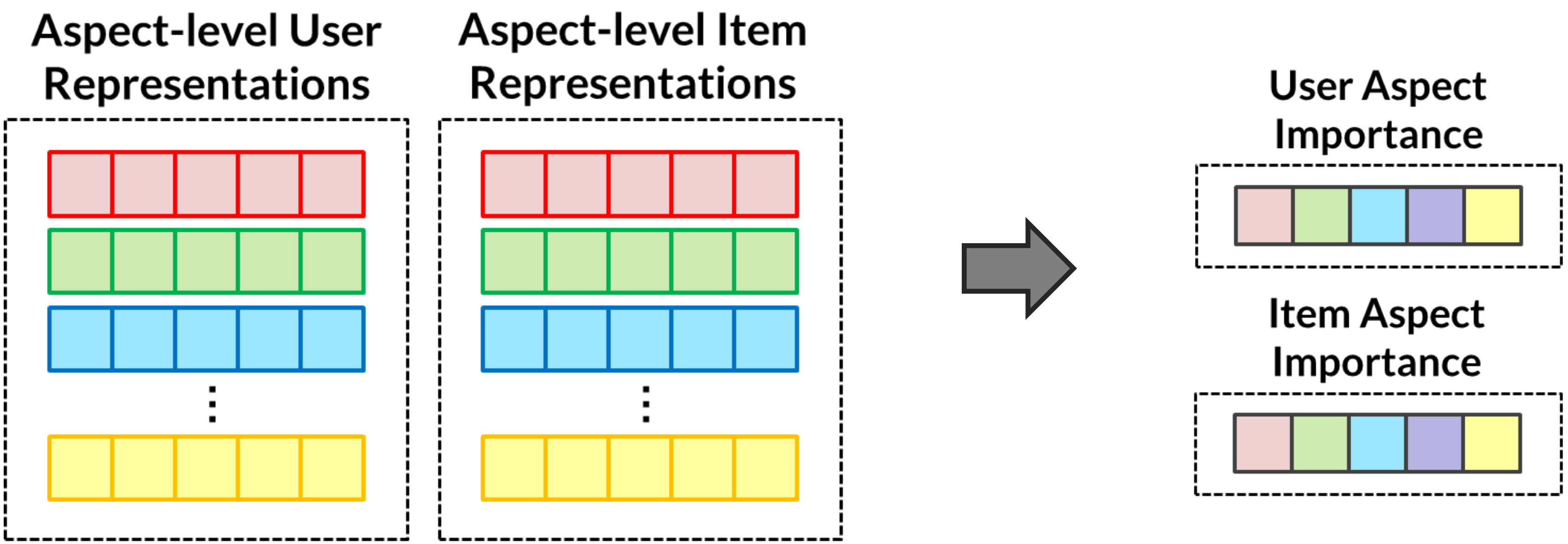
Item Aspect Importance:

$$\boldsymbol{H}_i = \phi\left(\boldsymbol{Q}_i \boldsymbol{W}_y + \boldsymbol{S}\,(\boldsymbol{P}_u \boldsymbol{W}_x)\right)$$

$$\boldsymbol{\beta}_i = softmax\left(\boldsymbol{H}_i \boldsymbol{v}_y\right)$$



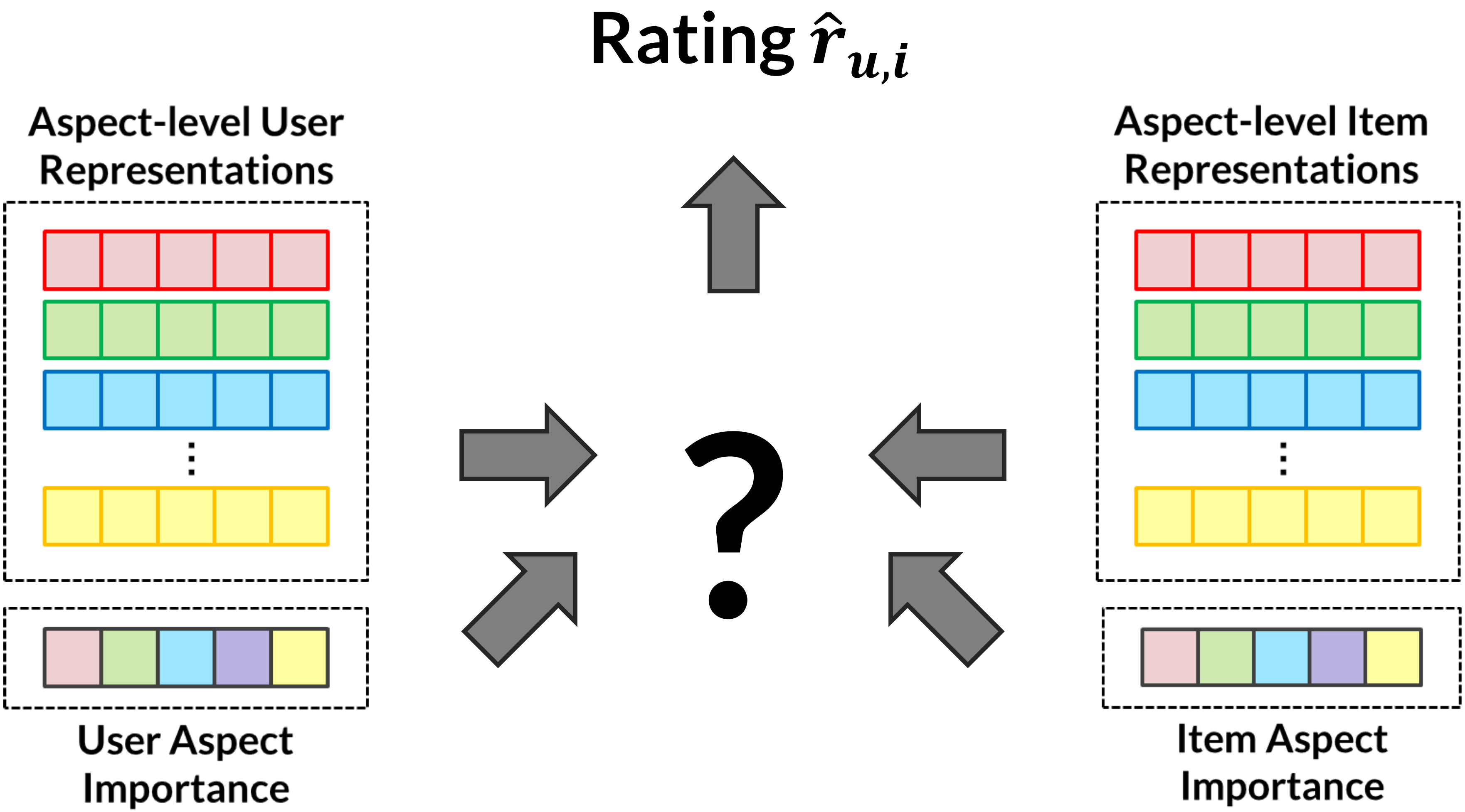
User & Item Aspect Importance



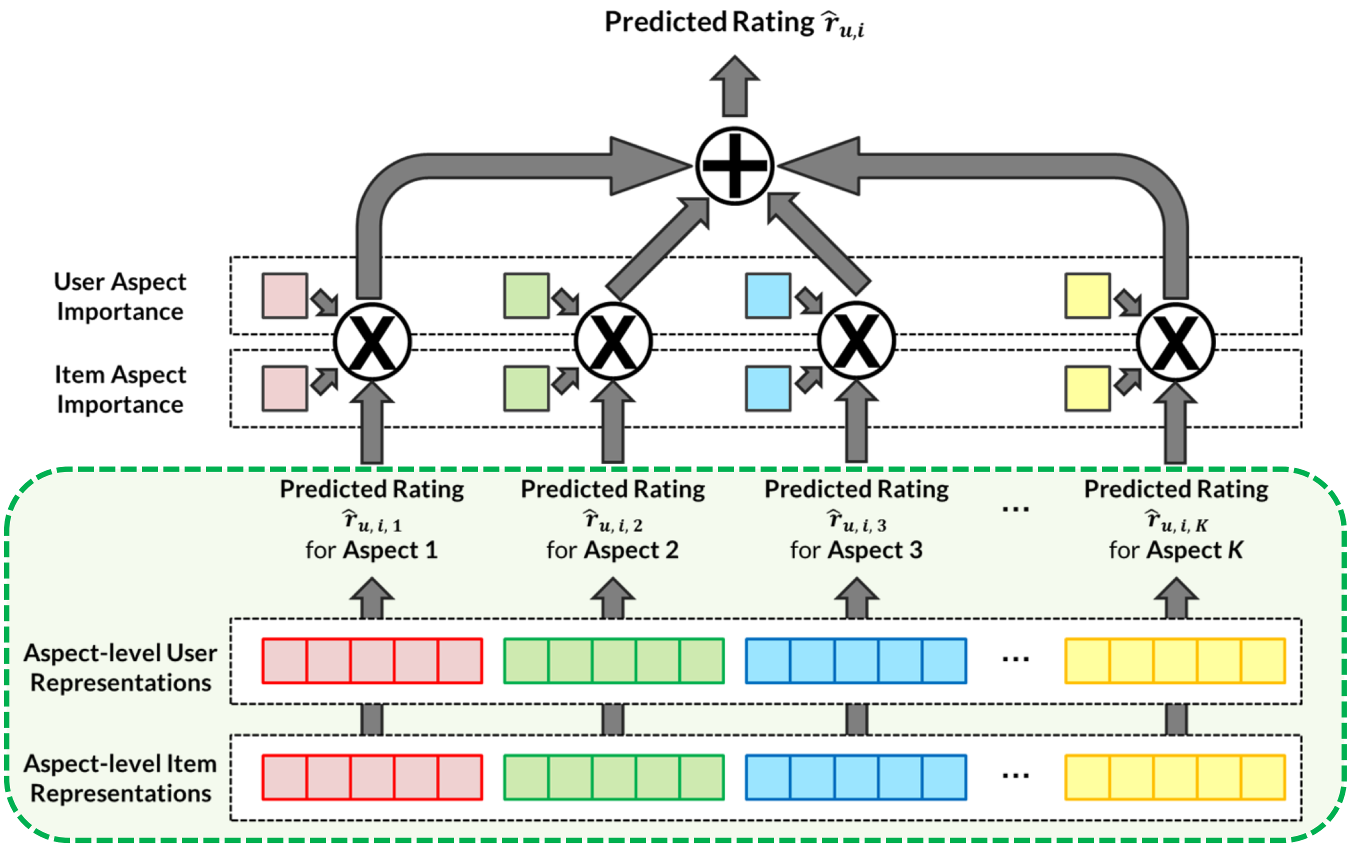
User & Item Aspect Importance are **interaction-specific** ☺

- ▷ User representations are used as the **context** for estimating item aspect importance, and vice versa
- ▷ Specifically **tailored** to each user-item pair

User-Item Rating Prediction



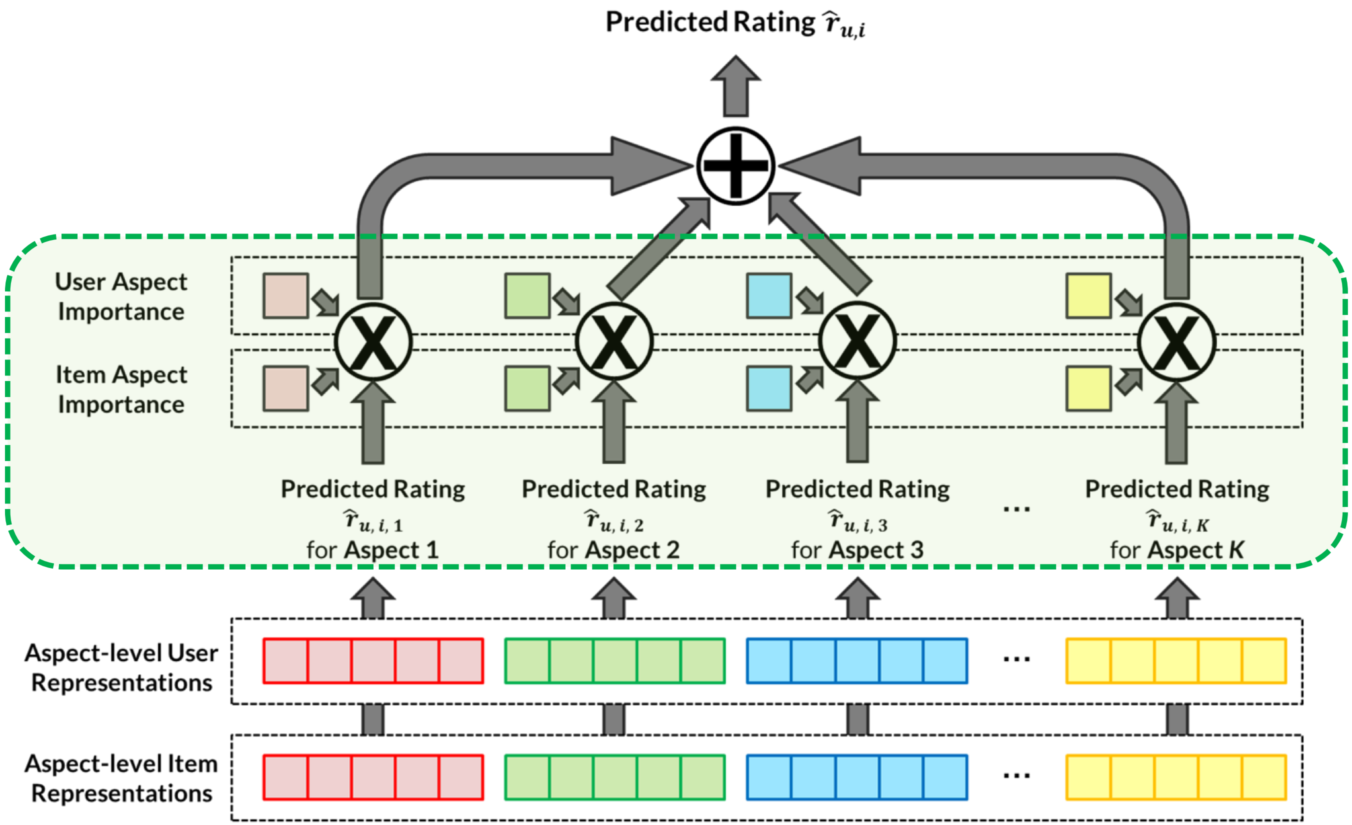
User-Item Rating Prediction



$$\hat{r}_{u,i} = \sum_{a \in \mathcal{A}} \left(\beta_{u,a} \cdot \beta_{i,a} \cdot \underbrace{\left(\boldsymbol{p}_{u,a} \left(\boldsymbol{q}_{i,a} \right)^{\mathsf{T}} \right)} \right) + b_u + b_i + b_0$$

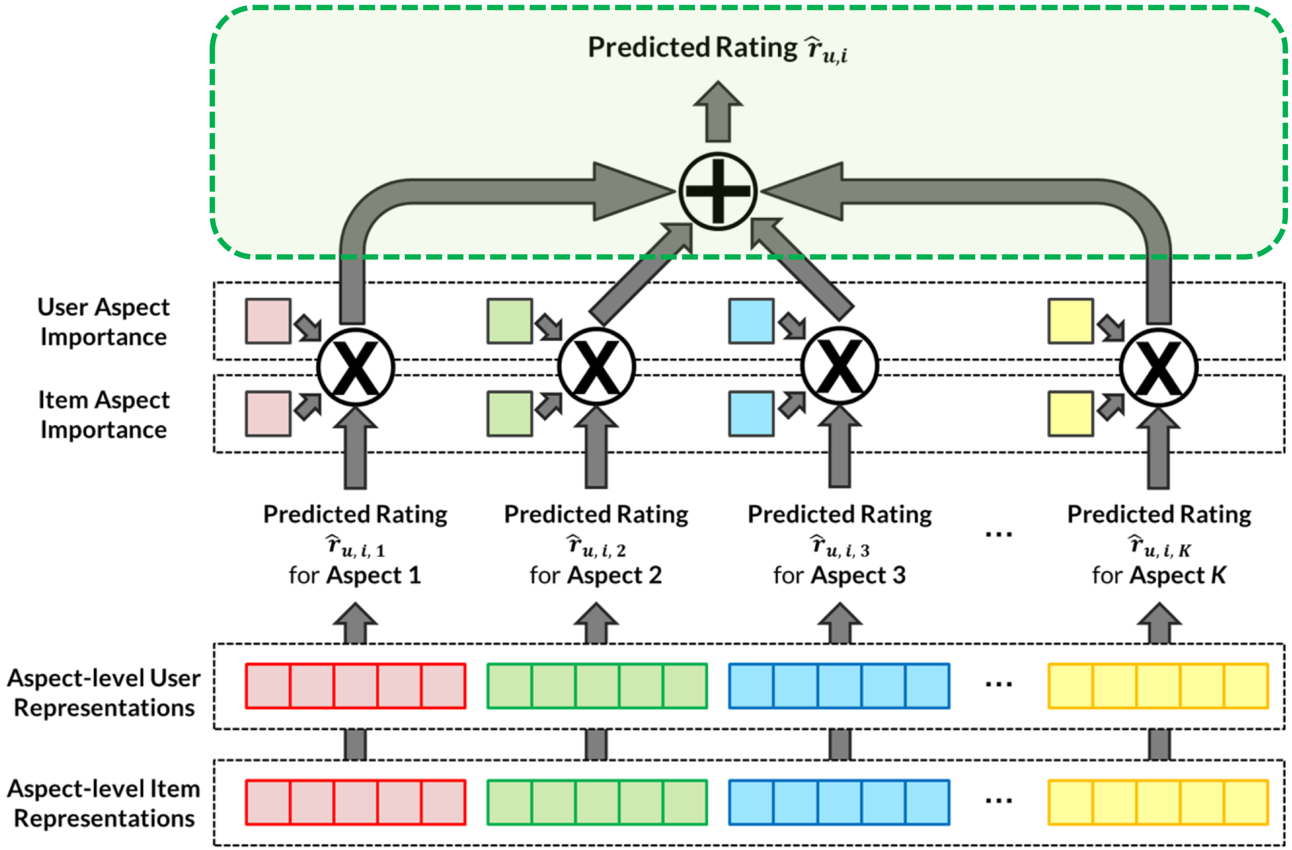
(1) Aspect-level representations → Aspect-level rating

User-Item Rating Prediction



$$\hat{r}_{u,i} = \sum_{a \in \mathcal{A}} \left(\underbrace{\beta_{u,a} \cdot \beta_{i,a} \cdot \left(\boldsymbol{p}_{u,a} \left(\boldsymbol{q}_{i,a} \right)^{\mathsf{T}} \right)}_{\text{(2) Weight by aspect-level importance}} \right) + b_u + b_i + b_0$$

User-Item Rating Prediction



$$\hat{r}_{u,i} = \underbrace{\sum_{a \in \mathcal{A}}}_{\text{(3) Sum across all aspects}} \left(\beta_{u,a} \cdot \beta_{i,a} \cdot \left(\boldsymbol{p}_{u,a} \left(\boldsymbol{q}_{i,a} \right)^{\top} \right) \right) + \underbrace{b_u + b_i + b_0}_{\text{(4) Include biases}}$$

Model Optimization

The model optimization process can be viewed as a regression problem.

- ▷ All model parameters can be learned using the backpropagation technique
- ▷ We use the standard Mean Squared Error (**MSE**) between the actual rating $r_{u,i}$ and the predicted rating $\hat{r}_{u,i}$ as the loss function

- ▷ Dropout is applied to each of the aspect-level representations
- ▷ L_2 regularization is used for the user and item biases

- ▷ Please refer to our paper for more details!

3.

Experiments & Results

Datasets

We use publicly available datasets from *Yelp* and *Amazon*

- **Yelp**
  - Latest version (Round 11) of the Yelp Dataset Challenge
  - Obtained from: https://www.yelp.com/dataset/challenge
- **Amazon**
  - Amazon Product Reviews, which has been organized into 24 individual product categories
  - For the larger categories, we randomly sub-sampled 5,000,000 user-item interactions for the experiments
  - Obtained from: http://jmcauley.ucsd.edu/data/amazon/
- For each of these **25 datasets**, we randomly select **80%** for training, **10%** for validation, and **10%** for testing

Baselines & Evaluation Metric

1. **Deep Cooperative Neural Networks (DeepCoNN), WSDM 2017**
   - Uses a convolutional architecture for representation learning, and performs rating prediction using a Factorization Machine
2. **Dual Attention-based Model (D-Attn), RecSys 2017**
   - Incorporates local and global attention-based modules prior to the convolutional layer for representation learning
3. **Aspect-aware Latent Factor Model (ALFM), WWW 2018**
   - Aspects are learned using an Aspect-aware Topic Model (**ATM**), and combined with a latent factor model for rating prediction

▷ **Evaluation Metric**

- Mean Squared Error (**MSE**) between the **actual rating** $r_{u,i}$ and the **predicted rating** $\hat{r}_{u,i}$

Experimental Results

Dataset	D-Attn (a)	DeepCoNN (b)	ALFM (c)	ANR (d)	Improvement (%) (d) vs. (a)	(d) vs. (b)	(d) vs. (c)
Amazon Instant Video	1.213	1.178	1.075	**1.009**	16.83	14.36	6.13
Apps for Android	1.637	1.593	1.555	**1.412**	13.73	11.34	9.14
Automotive	1.411	1.349	1.257	**1.188**	15.76	11.91	5.43
Baby	1.507	1.442	1.359	**1.258**	16.51	12.73	7.44
Beauty	1.609	1.566	1.466	**1.386**	13.89	11.48	5.46
Books	1.122	1.089	1.055	**0.976**	12.94	10.30	7.43
CDs & Vinyl	1.014	0.980	0.956	**0.914**	9.93	6.81	4.46
Cell Phones & Accessories	2.083	2.040	1.787	**1.689**	18.92	17.23	5.50
Clothing, Shoes & Jewelry	1.491	1.430	1.316	**1.266**	15.09	11.48	3.78
Digital Music	0.775	0.749	0.725	**0.688**	11.22	8.12	5.07
Electronics	1.744	1.659	1.563	**1.445**	17.10	12.89	7.50
Grocery & Gourmet Food	1.386	1.345	1.284	**1.187**	14.42	11.76	7.57
Health & Personal Care	1.612	1.545	1.466	**1.356**	15.91	12.23	7.49
Home & Kitchen	1.575	1.508	1.443	**1.317**	16.38	12.69	8.76
Kindle Store	0.949	0.905	0.870	**0.834**	12.08	7.81	4.10
Movies & TV	1.246	1.207	1.193	**1.112**	10.75	7.88	6.80
Musical Instruments	1.224	1.160	1.072	**1.034**	15.51	10.81	3.49
Office Products	1.650	1.569	1.474	**1.337**	18.98	14.79	9.30
Patio, Lawn & Garden	1.696	1.622	1.510	**1.403**	17.30	13.51	7.09
Pet Supplies	1.628	1.565	1.485	**1.377**	15.41	12.05	7.28
Sports & Outdoors	1.354	1.300	1.221	**1.137**	16.04	12.55	6.86
Tools & Home Improvement	1.474	1.429	1.348	**1.230**	16.51	13.93	8.74
Toys & Games	1.298	1.227	1.131	**1.075**	17.16	12.34	4.88
Video Games	1.533	1.498	1.383	**1.292**	15.72	13.72	6.57
Yelp (2018)	1.691	1.669	1.614	**1.527**	9.68	8.49	5.42
Average	1.437	1.385	1.304	**1.218**	14.95	11.73	6.47

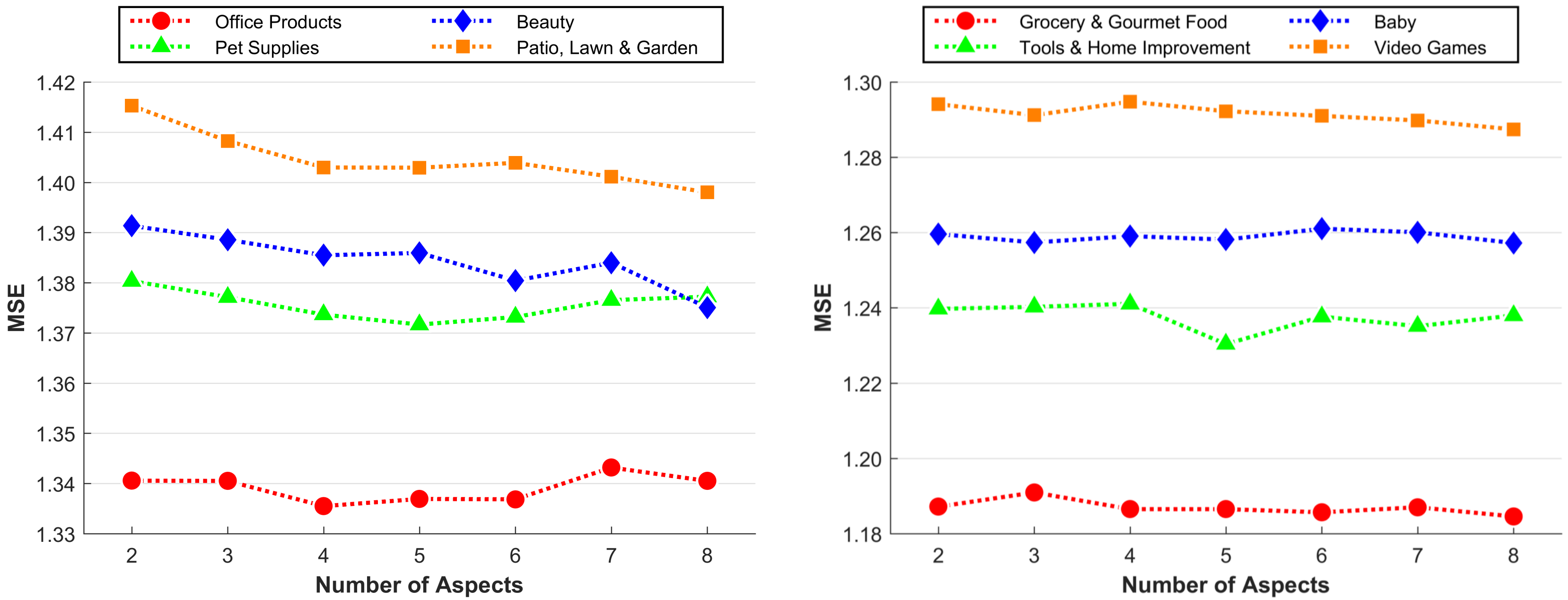
Experimental Results

- ▷ **Statistically significant improvements** over all 3 state-of-the-art baseline methods, based on the *paired sample t-test*
  - The average improvement over D-Attn, DeepCoNN, and ALFM are **14.95%**, **11.73%**, and **6.47%**, respectively

- ▷ Outperforms D-Attn and DeepCoNN due to 2 main reasons:
  - Instead of having a single 'compressed' user and item representation, we learn **multiple aspect-level representations**
  - Additionally, we estimate the **importance of each aspect**

- ▷ We outperform a similar aspect-based method ALFM as we learn both the aspect-level representations and importance in a **joint manner**

Number of Aspects



- ▷ Key Hyperparameter: Number of Aspects
- ▷ In our experiments, we use **5 aspects** to be consistent with ALFM
- ▷ **Relatively stable performance** for a reasonable number of aspects
  - A handful of broader aspects
  - Numerous fine-grained aspects

Learned Aspects

Price	Family	Negative	Gameplay	Graphics
works	son	bad	lot	bought
recommend	new	little	hours	pretty
well	highly	horrible	bit	still
buy	story	waste	couple	graphics
bought	favorite	hard	characters	much
awesome	part	boring	stars	think
price	character	terrible	course	work
loves	daughter	frustrating	minutes	recommend
worth	controller	difficult	side	cool
purchase	characters	disappointed	fan	nice

- Aspects are learned in a **data-driven manner** without any external supervision
- We use the **words** with the **highest attention scores** (averaged across all users & items) to represent each aspect

4.

Future Work & Conclusion

Future Work

1. Explainable Recommendation

- ▷ For each user-item interaction, ANR is capable of estimating the **importance of each aspect**
- ▷ For the top K (most important) aspects, we can identify the **relevant document segments** which contribute to its representation

2. Domain-independent Aspect-based Recommendation

- ▷ Currently, a **separate model** needs to be trained for each category/domain
- ▷ Extend ANR into a **domain-independent framework**, which will be able to handle multiple categories simultaneously, by incorporating either **transfer learning** or **multi-task learning**

Summary

- We proposed an *Aspect-based **N**eural **R**ecommender* (**ANR**) to leverage the strengths of both deep learning techniques and aspect-based recommender systems

- **Aspect-level representations** are learned by focusing on relevant words in the document using the ***neural attention mechanism***
- **Interaction-specific aspect importance** are estimated using the user and item aspect-level representations by extending the ***neural co-attention mechanism***
- We effectively combine the aspect-level representations and importance to derive the **aspect-level ratings**, which are used for estimating the overall rating

Thanks!

Any questions?

Email:
S160005@e.ntu.edu.sg