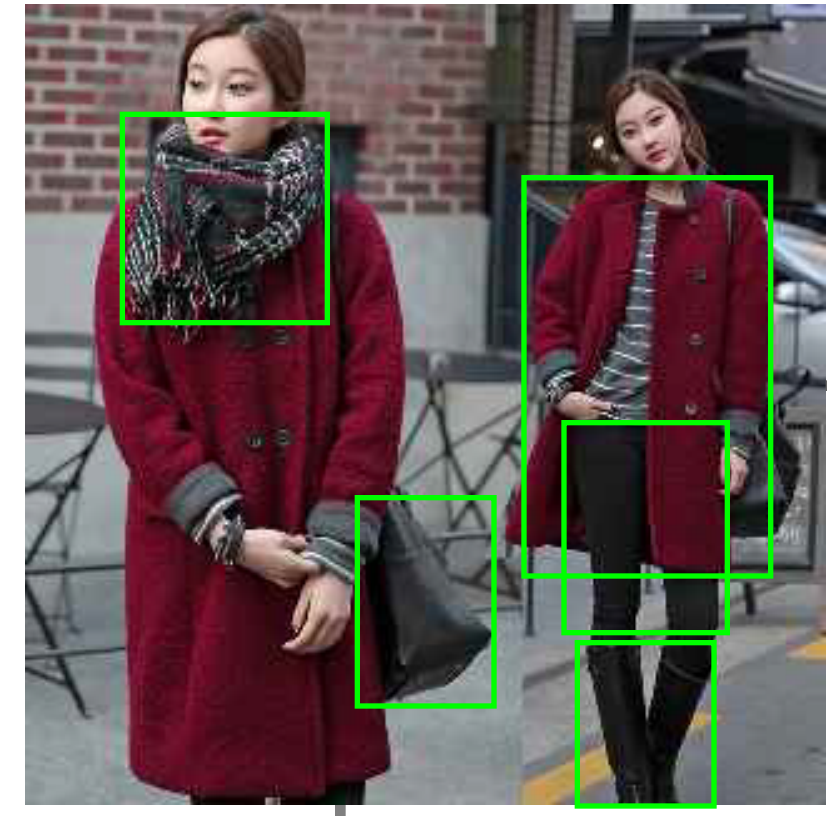


MSE Extraction Layers

Symmetric Linear GCN Layers

Prediction & Optimization



Image

Pre-trained
PNASNet

output

get top- k
classes

red coat
boots
scarf
handbag
jeans

Visual semantic entity mining

 i Item-entity Affiliation
Graph \mathcal{G}_2

Title and reviews

Title: Red cashmere female
winter mid-length coat.

Review 1: I received the **coat**, its
quality and the **shape** are very
good, the **upper body** looks good
and **thin**, did not let me down.

...

Review x : The clothes are very
thick and **warm** to wear, and
most importantly, they are also
very **cheap**.

Keyword
Extraction

female good
winter red
cashmere
mid-length
coat cheap
thin quality
good shape
thick warm
upper body

Textual semantic entity mining

Sentiment
Analysis

mean

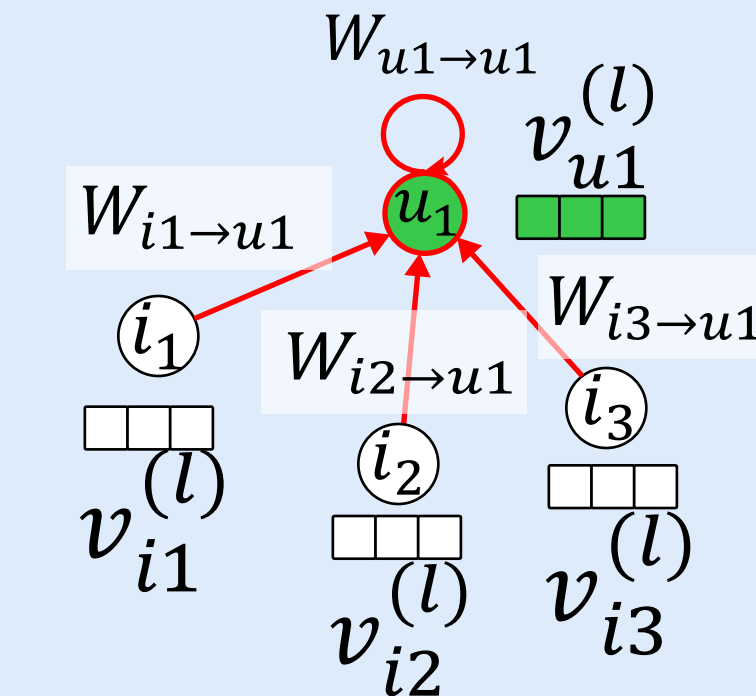
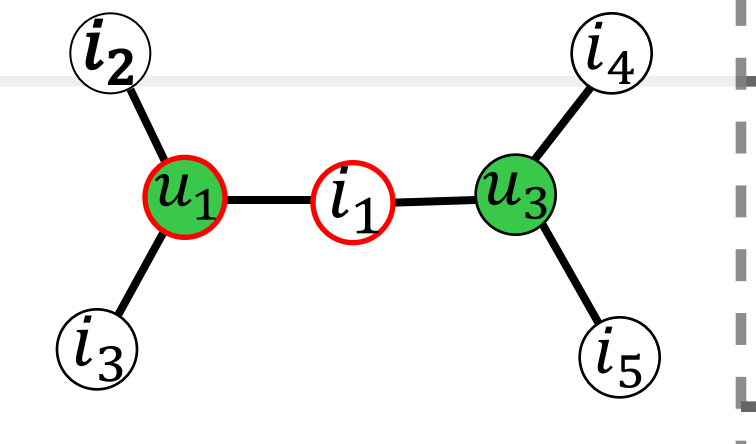
 S_i

normalize

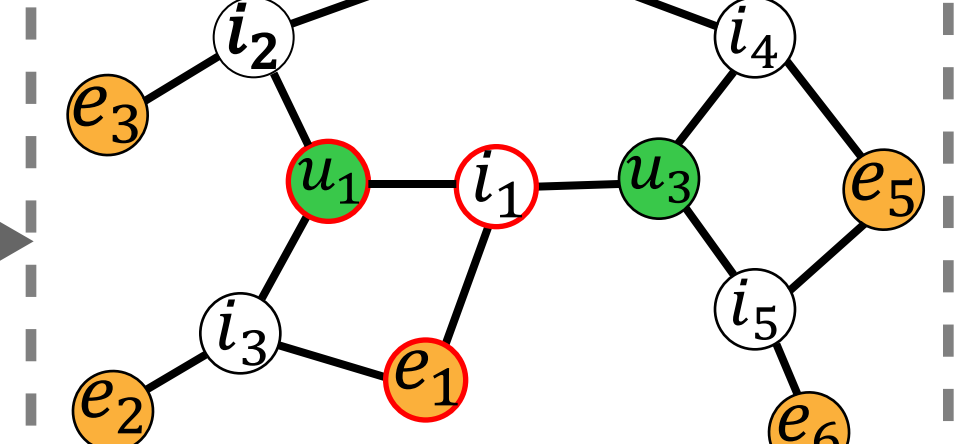
$$\frac{(s^i)^{0.1} |I|}{\sum_{i \in I} (s^i)^{0.1}}$$

Assign weights to all item nodes

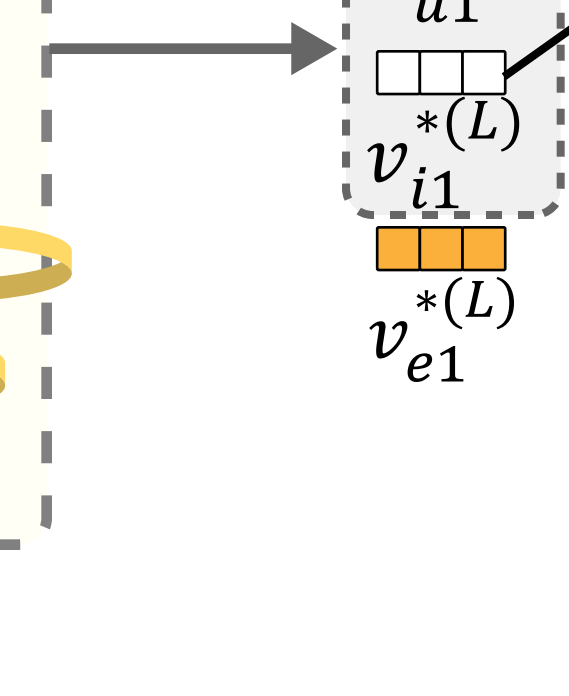
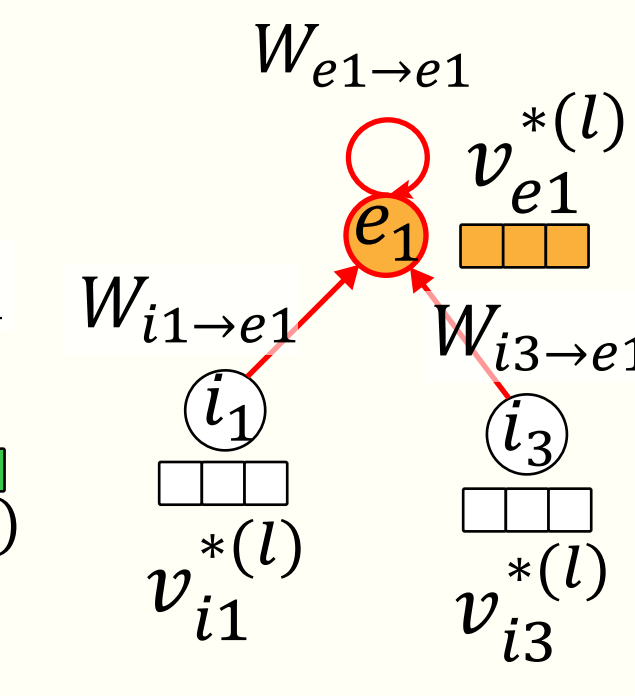
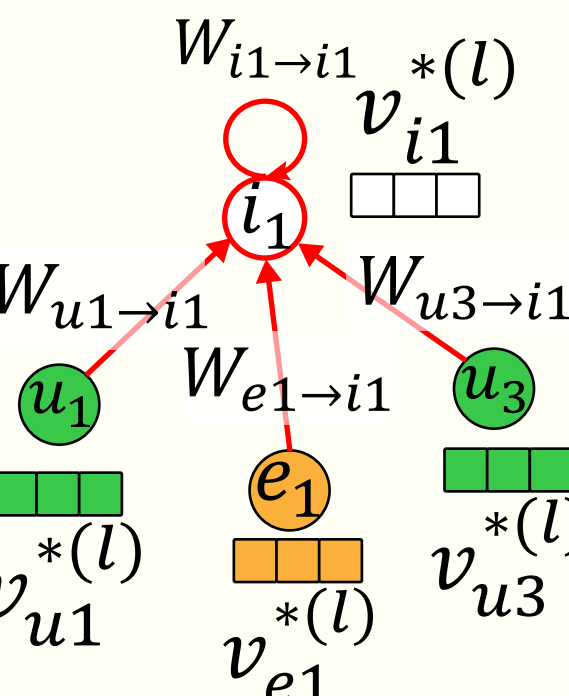
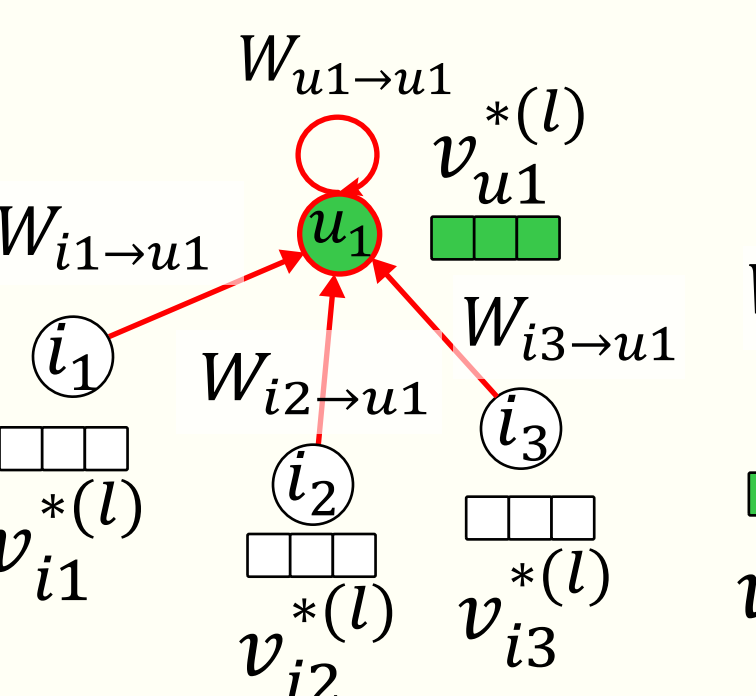
LS-GCN-1

User-item Interaction
Graph \mathcal{G}_1 

fuse

Collaborative Multimodal
Interaction Graph \mathcal{G} 

LS-GCN-2

optimize \mathcal{L}_2 optimize \mathcal{L}_1 \oplus \hat{y}_{u1i1} \otimes inner product \oplus addition u user node i item node e MSE node

$$W_{u \rightarrow i} = \frac{(s_i)^{0.1} \cdot |I|}{\sum_{i \in I} (s_i)^{0.1}} \cdot \frac{1}{|N_u|^{0.5} \cdot |N_i|^{0.5-\alpha}}$$