

Raw Input

Review Text

$R_{u_j}^{(1)}$

$R_{u_j}^{(k)}$

$R_{u_j}^{(t)}$

Multiple review texts at t timestamps for user j

Pretrained Embedding

Review Text Module

$s_u^{(k)} \in \mathbb{R}^{l_r * d_e}$

Convolution

$E_u^k \in \mathbb{R}^{l_d * l_r * d_c}$

Word-Level
Multi-Dim Attention

Temporal Preference
Evolution

T_u^1

T_u^k

T_u^t

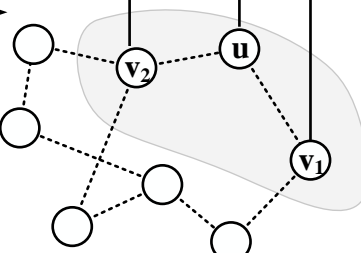
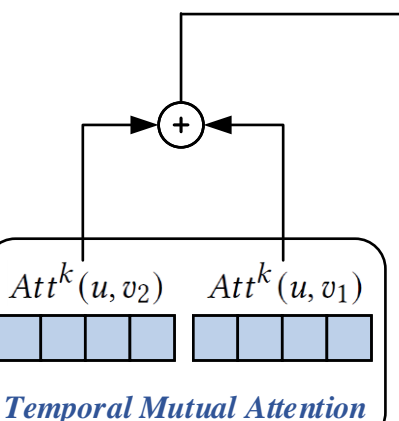
h_u^1

h_u^k

h_u^t

GRU

Social Attention Module



Social Network

Social Attention
Evolution

M_u^1

M_u^k

M_u^t

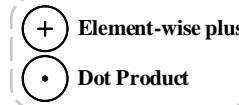
g_u^1

g_u^k

g_u^t

GRU

Rating Prediction



Ground-Truth Label

$r_{u,i}$

Loss

$\hat{r}_{u,i}$

Multi-layer
Decoder

