

## **Selection and Generation: Learning towards Multi-Product Advertisement Post Generation**

**Zhangming Chan<sup>1,2,\*</sup>, Yuchi Zhang<sup>3</sup>, Xiuying Chen<sup>1,2</sup>, Shen Gao<sup>2</sup>,  
Zhiqiang Zhang<sup>3</sup>, Dongyan Zhao<sup>1,2</sup> and Rui Yan<sup>1,2,†</sup>**

<sup>1</sup> Center for Data Science, AAIS, Peking University, Beijing, China

<sup>2</sup> Wangxuan Institute of Computer Technology, Peking University, Beijing, China

<sup>3</sup> Alibaba Group, Beijing, China

{zhangming.chan, xy-chen, shengao, zhaody, ruiyan}@pku.edu.cn

{yuchi.zyc, zhang.zhiqiang}@alibaba-inc.com

# Motivation

- A good AD post can highlight the characteristics of each product, thus helps customers make a good choice among candidate products
- all the previous studies focus on copywriting generation for a single product, consumers need to compare among products and summarize the advantages and disadvantages of each product by themselves.

# Motivation

- A multi-product AD post contains several related products that either function similarly or match each other. Each product has its own copywriting
- Each product's copywriting will take full account of the topic of the post and other product's information; meanwhile, it also contains its unique features.
- The multi-product AD post writing is more difficult, first , it needs to **select suitable products** for writing in a post; to describe the unique characteristics of each product, the copywriter should also **consider the information of other products in the post**

# Motivation

## 6款网红麦片, 哪个好吃又不胖

6款产品

一向注重饮食的超模们, 早上都会毫不犹豫来一顿麦片餐, 比如刘雯。各种美食博主, 都爱 Po 燕...



¥82



已售罄



已售罄

if姐

2.2W 2

### 家乐氏玫瑰麦片

这款开袋就有浓郁的玫瑰香气扑鼻而来, 干吃的话很合适, 一天一袋, 吃起来香甜酥脆。当早餐配牛奶、配酸奶也可以, 泡发之后能增加玫瑰的香气, 里面有椰蓉, 白白的、细细的、小小的, 是清新的椰子味。不过它的热量和含糖量很高, 想减肥的宝宝不要入手啊。

### ICA什锦燕麦片

这款麦片里面除了酸奶球, 还有酸甜的草莓干、脆脆的玉米片等等, 量大而足。泡牛奶也非常 nice, 麦片很脆, 牛奶变得粉粉的, 这不是草莓干掉色, 而是草莓粉末溶在了牛奶里。整体吃起来, 带着草莓的酸甜, 如果不小心吃到了酸奶片, 还有种隐秘的小惊喜感。

Six oatmeals which are popular now, let's find the one is yummy but will not let you be fat.

### Kellogg's Tasty Granola-Rose

You'll be showered with a rich fragrant of roses when you tear open a bag of this granola. It is a great snack on its own, and with its crispy, sweet taste, you'll find it easy to finish a whole bag in one day. It also makes a delicious breakfast if you eat it with milk or yogurt, which will bring out the sweet smell of the roses even more. The granola also contains fine desiccated coconut that adds a refreshing coconut flavor to the taste. However, this type of granola is high in both calories and sugar, so it's probably not the best choice for those of you who are on a diet.

### ICA oatmeal-crunchy jordgubbar & yogurt

This oatmeal contains not only a satisfying amount of yogurt balls, but also sour-sweet dried strawberries and crispy corn flakes. The crispy cereal goes great with milk, which turns into a pretty pink as powders of strawberries dissolve. (Don't worry, the strawberries are not dyed and do not bleed). Overall, it has a sour-sweat flavor that comes from the strawberries, and the yogurt flakes that occasionally pop into a bite will make a nice little surprise,

# Problem Formulation

- Use  $P$  to denote a product candidate set which contains a lot of productions, namely  $P = \{u_1^p, u_2^p, \dots, u_n^p\}$ , where  $u_i^p$  indicates the information of  $i$ -th product in the set. Specifically,  $u_i^p = \{w_{i,1}^p, w_{i,2}^p, \dots, w_{i,n_{p,i}}^p\}$  is a text sequence that contains  $n_{p,i}$  words.  $T = \{w_1^t, w_2^t, \dots, w_{n_t}^t\}$  is the topic sentence. The goal of our model is to generate a multi-product AD-post  $\hat{C} = \{\hat{u}_1^c, \hat{u}_2^c, \dots, \hat{u}_n^c\}$  where  $\hat{u}_i^c$  represents the copywriting of the  $i$ -th product in post.

# Method

- In this model, we split the multi-product AD post generation task into two sub processes: (1) select a set of products via the SelectNet (Selection Network). (2) generate a post including selected products via the MGenNet (Multi-Generator Network). It generate each product's copywriting considering the topic and other products

# Method

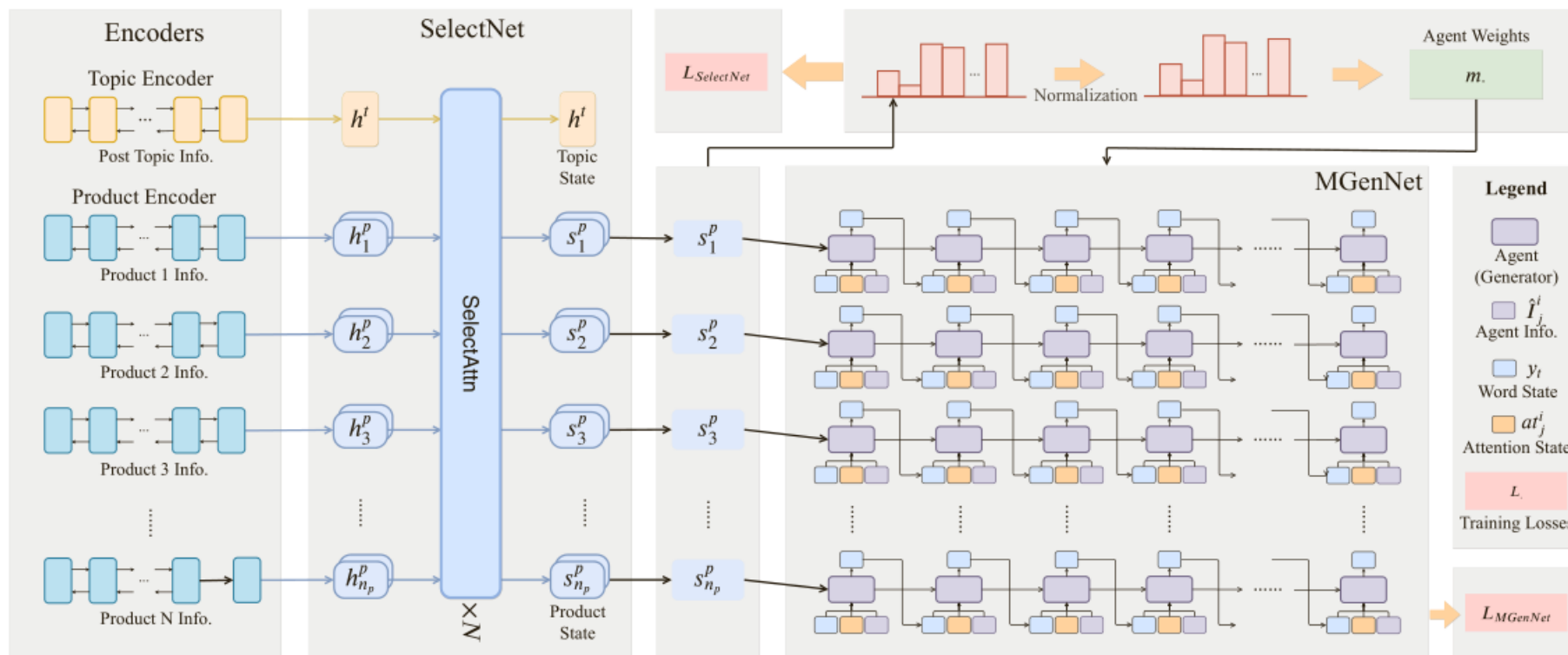


Figure 3: The overview of the end-to-end S-MG Net which is the combination of SelectNet and MGenNet model.

# Selection Network

- Given the post topic  $T$  and the product candidate set  $P$
- we use an RNN encoder named topic encoder to encode the embedded post topic  $e(T)$  to  $h^t$
- Similarly, we use another RNN encoder called product encoder to encode each embedded product information sequence  $e(u_i^p)$  as  $h_i^p$ . So we can obtain  $h^P = \{h_1^P, h_2^P, \dots, h_{n_p}^P\}$  which indicates all products.



# Selection Network

- To capture the relationship among the product candidates, we use the self-attention mechanism to conduct the interaction between  $h^P = \{h_1^P, h_2^P, \dots, h_{n_p}^P\}$ . We use three fully-connected layer to project  $h_i^p$  into three spaces:  $Q_i = FC_q(h_i^p)$ ,  $K_i = FC_k(h_i^p)$ ,  $V_i = FC_v(h_i^p)$ .

$$\alpha_{i,j} = \frac{\exp(Q_i K_j)}{\sum_{n=1}^{n_p} \exp(Q_i K_n)},$$
$$\beta_i = \sum_{j=1}^{n_p} \alpha_{i,j} v_j,$$

$$\hat{s}_i^p = \text{LayerNorm}(h_i^p + \beta_i),$$

- We then obtain a new representation  $\widehat{s}_i^p$ :

$$\hat{s}_i^p = \text{LayerNorm}(h_i^p + \beta_i),$$

# Selection Network

- Next, we apply a feed-forward layer on  $\widehat{s}_i^p$  to obtain  $\dot{s}_i^p$  and conduct the residual connection layer again:

$$\begin{aligned}\dot{s}_i^p &= \text{ReLU}(\hat{s}_i^p \cdot W_1 + b_1) \cdot W_2 + b_2, \\ \dot{s}_i^p &= \text{LayerNorm}(\dot{s}_i^p + \hat{s}_i^p),\end{aligned}$$

- to target at utilizing the relationship between post topic and each product, we propose to integrate the information of post topic into each product presentation  $\dot{s}_i^p$ :

$$\begin{aligned}g^p &= \text{Sigmoid}([\dot{s}^p; h^t] \cdot W_g + b_g) \\ \widetilde{s}^p &= g^p \cdot h^t + (1 - g^p) \cdot \dot{s}^p, \\ s^p &= \dot{s}^p + \widetilde{s}^p,\end{aligned}$$

# Selection Network

- After adding post topic information to product representation, we use the raw product information to polish the topic representation

$$\gamma_i = \frac{\exp(\dot{s}_i^p h^t)}{\sum_{j=1}^{n_p} \exp(\dot{s}_j^p h^t)},$$
$$\delta = \sum_{i=1}^{n_p} \gamma_i \dot{s}_i^p.$$

- Inspired by GLU, we use the product and topic information to control the amount of information of  $\delta$ .

$$g^t = \text{Sigmoid}([\delta; h^t] \cdot W_t + b_t),$$
$$\tilde{\delta} = g^t \cdot \delta,$$
$$h^t = h^t + \tilde{\delta}$$

# Selection Network

- We name the whole above operation as SelectAttn. We finally obtain new product representations  $s^P = \{s_1^P, s_2^P, \dots, s_{n_p}^P\}$
- After interacting with the post topic and other products, the product representations can describe the matching degree of the product with the target post:

$$score. = \text{Sigmoid}(s^P \cdot W_c),$$

# Multi-Generator Network

- As mentioned before, we rank the selection scores  $\{score_1, score_2, \dots, score_{n_p}\}$  to get top M products as selection result  $\overline{s^p} = \{\overline{s_1^p}, \overline{s_2^p}, \dots, \overline{s_M^p}\}$ . Where  $\{\widehat{m_1}, \widehat{m_2}, \dots, \widehat{m_M}\}$  . is the corresponding scores. We normalize these scores as:

$$m_{.} = \frac{\widehat{m_{.}} - \text{Min}(\widehat{m_{.}})}{\text{Max}(\widehat{m_{.}}) - \text{Min}(\widehat{m_{.}})},$$

- We use a linear layer to cover all the selected product representation  $\{\overline{s_1^p}, \overline{s_2^p}, \dots, \overline{s_M^p}\}$  to initialize the corresponding agent (RNN Cell):

$$\begin{aligned} a_0 &= \text{ReLU}(\overline{s^p} \cdot W_s) + b_s, \\ a_1 &= \text{Agent}(a_0, e(<\text{BOS}>)), \end{aligned}$$

# Multi-Generator Network

- Let  $a_i$  represents the state of agent in the i-th step. During the generation, we let the agents communicate with each other.

$$\begin{aligned}\hat{a}_j &= a_j \cdot m., \\ I_j^i &= \{\hat{a}_j^1, \dots, \hat{a}_j^{i-1}, \hat{a}_j^{i+1}, \dots, \hat{a}_j^M\}\end{aligned}$$

- Where  $\hat{a}_j^i$  represents the information coming from the i-th agent when j-th step. We use the mean value of  $I_j^i$  as the practical information for the i-th agent in the j-th step as

$$\hat{I}_j^i = \frac{\text{sum}\{I_j^i\}}{M - 1}.$$

$$\begin{aligned}y'_t &= ([y_t; \hat{I}_j^i; at_t] \cdot W_y) + b_y, \\ a_{t+1} &= \text{Agent}(a_t, y'_t),\end{aligned}$$

# Training objective

- We launch the following objective to minimize the MLE loss between the ground truth and generated copywriting. Meanwhile, we minimize the loss between selected product ground truth in real multi-product AD post

$$\mathbb{L} = \lambda \mathbf{E}_{\theta} \log p(\hat{P}|P, T) + \gamma \mathbf{E}_{\phi} \log p(\hat{C}|\hat{P}, T),$$

# Experiment

We construct a multi-product AD post dataset collected from Bimai Qingdan (which means a list of goods you must buy) in Taobao. Millions of posts are composed by professional copywriters to introduce and recommend different products for online shoppers. Each post consists of several products with their images, description copywriting, and a title that reflects the topic of this post.



# Experiment

Models	Embedding Metrics			Inter-Distinct				Intra-Distinct			
	Average	Greedy	Extrema	Dist-1	Dist-2	Dist-3	Dist-4	Dist-1	Dist-2	Dist-3	Dist-4
Seq2seq	0.9197	548.69	0.4293	0.937	3.183	6.655	9.940	20.77	28.38	31.07	33.41
ConvSeq	0.6049	326.99	0.1123	1.308	2.796	3.970	4.951	-	-	-	-
Transformer	0.8662	537.69	0.3941	1.473	4.427	9.099	13.59	25.23	33.83	39.79	45.48
PCPG	0.8830	540.41	0.3713	1.409	3.943	7.423	10.43	22.43	29.98	36.80	40.31
KOBE	0.8783	539.23	0.4023	1.523	5.334	11.34	18.32	26.46	37.43	43.23	53.84
<b>S-MG<sub>C</sub></b>	<b>0.9438</b>	560.45	<b>0.4481</b>	<b>1.763</b>	<b>8.051</b>	<b>18.37</b>	<b>28.30</b>	<b>44.66</b>	<b>66.22</b>	<b>73.57</b>	<b>78.96</b>
<b>S-SG</b>	0.8774	<b>566.86</b>	0.4280	1.294	4.059	8.479	12.71	24.76	33.06	38.72	44.09
<b>S-MG</b>	0.9428	558.62	0.4440	1.713	7.502	17.21	26.60	44.49	65.97	73.26	78.64

Title	赶快解锁女装利器，从此迎接美好生活(Unlock the weapon of women's clothes, and welcome the beautiful life from now on)
Product Set	<p>(×) 1. d'zzit 地素夏专柜新款运动风半身裙(D'zzit's new sporty style skirt which is designed for the summer.)</p> <p>(✓) 2. vifille 同款刺绣撞色卫衣 (This hoodie has vibrant color and embroidery and it is so fashione that Vifille also wear it !)</p> <p>(✓) 3. beanpole 滨波女士宽松版印花卫衣(BEAN POLE's women hoodie with printing and it is loose for wearing.)</p> <p>(×) 4. 2016 夏季新款女装名媛条纹不规则假两件开叉露肩连衣裙夏中长款 (Fake two off-the-shoulder split women dresses with irregular stripes, and it is the special style in 2016)</p> <p>(×) 5. 优雅七分袖夏季宽松a字裙(It is an elegant summer A-line skirt with three-quarter sleeves.)</p> <p>(×) 6. 白夜宽松蝙蝠袖套头毛衣(White-night's bat-sleeved sweater, and it is loose for wearing.)</p> <p>(✓) 7. 果酱公主春新原创森系时尚气质卫衣 (Guojiang Princess's original fashion hoodie for the spring and you will be very comfortable because it is like fresh air.)</p> <p>(×) 8. 梦舒雅冬新款加厚加绒微喇裤修身(Mengshuaya's new thickened pants and it has fleecy trim.)</p> <p>(✓) 9. 七格格卫衣加绒半高领秋中长款宽松 (Qigege's hoodie with fleece trim and semi-high collar. It is the middle length for autumn. It is loose for wearing.)</p> <p>(×) 10. 休闲七分显瘦通勤正装裤(7-points pants for relaxing, it will make you look more thin.)</p> <p>(×) 11. yiner 冬时尚贴布侧条纹系带连衣裙(Yiner's fashion dress with side stripes and straps, it is suitable for winter.)</p> <p>(×) 12. loftshine 珞炫女装专柜大衣(Loftshine's clothing coat, it is designed for women.)</p> <p>(×) 13. 洋气短袖碎花小清新超仙气质雪纺衫(This skirt which has short sleeve and the broken flower is pure and fresh.)</p> <p>(✓) 14. 卫衣女nina mist 套头飘带(This hoodie is designed for women by Nina Mist's and it has a beautiful ribbon.)</p> <p>(✓) 15. 韩版学生蝙蝠袖连帽卫衣(Korean style hoodie with bat sleeve. It is suitable for students.)</p>
S-MGc	<ul style="list-style-type: none"> <li>● <b>[For Product 14]:</b> 这款卫衣采用了经典的圆领设计，简约大方，宽松的版型 (This hoodie adopts the classic round collar design, it is simple and generous. Meanwhile, its version is loose.)</li> <li>● <b>[For Product 7]:</b> 这款卫衣采用了柔软的针织面料，上身舒适又保暖 (This hoodie is soft and like air, and it keeps the body warm and comfortable.)</li> <li>● <b>[For Product 15]:</b> 韩版风格，充满优雅气息，版型宽松，上身就是时髦 (The hoodie is Korean style and elegant. Its loose version makes owner fashion.)</li> <li>● <b>[For Product 9]:</b> 中长款的长度，百搭不挑身形，版型宽松，上身就是时髦 (The hoodie is middle length and it is suitable for multiple body shapes. Its loose version makes owner fashion.)</li> <li>● <b>[For Product 2]:</b> 精致的刺绣，点缀着整体的时尚不呆板，经典的圆领 (This clothes has delicate embroidery which shows the fashion, and it has classic round collar.)</li> <li>● <b>[For Product 3]:</b> 这是来自BEAN POLE 的卫衣，适合女士在休闲场合穿着 (This is a hoodie from BEAN POLE and it is suitable for casual wearing.)</li> </ul>

# **Learning a Simple and Effective Model for Multi-turn Response Generation with Auxiliary Tasks**

**Yufan Zhao<sup>1</sup>, Can Xu<sup>1\*</sup>, Wei Wu<sup>2</sup>**

<sup>1</sup>Microsoft Corporation, Beijing, China

<sup>2</sup>Meituan, Beijing, China

{yufzhao, caxu}@microsoft.com

wuwei19850318@gmail.com

# Motivation

- The existing state-of-the-art multi-turn response generation model addresses the problem with deep neural architectures
- While these models improved response quality, their complexity also hinders the application of the models in real systems.

# Motivation

- In this work, we pursue a model that has a simple structure yet can effectively leverage conversation contexts for response generation.
- To this end, we propose four auxiliary tasks including word order recovery, utterance order recovery, masked word recovery, and masked utterance recovery, and optimize the objectives of these tasks together with maximizing the likelihood of generation

# Problem Formulation

- We have a dataset  $D = \{(U_i, R_i)\}_{i=1}^N$ , where  $U_i = \{U_{i,1}, U_{i,2}, \dots, U_{i,n}\}$  denotes a context with  $U_{i,j}$  the  $j$ -th utterance.  $R_i$  is the response corresponding to  $U_i$ . A common practice is to learn  $P(R|U)$  by maximizing the log-likelihood

$$\sum_{i=1}^N \log P(R_i | \mathcal{U}_i).$$

# Generation Model

- We unfold all words in  $(U, R)$  into  $W = (w_1, \dots, w_m, w_{m+1}, \dots, w_{m+t})$ , where  $m$  is the number of words in context  $U$ . Suppose  $(w_{m+1}, \dots, w_{m+l-1})$  are words generated until step  $l - 1$ , the next word is predicted according to:

$$P(w_{m+l}|w_1, \dots, w_{m+l-1}) = \text{softmax}(W_s O(w_{m+l-1})),$$

# Generation Model

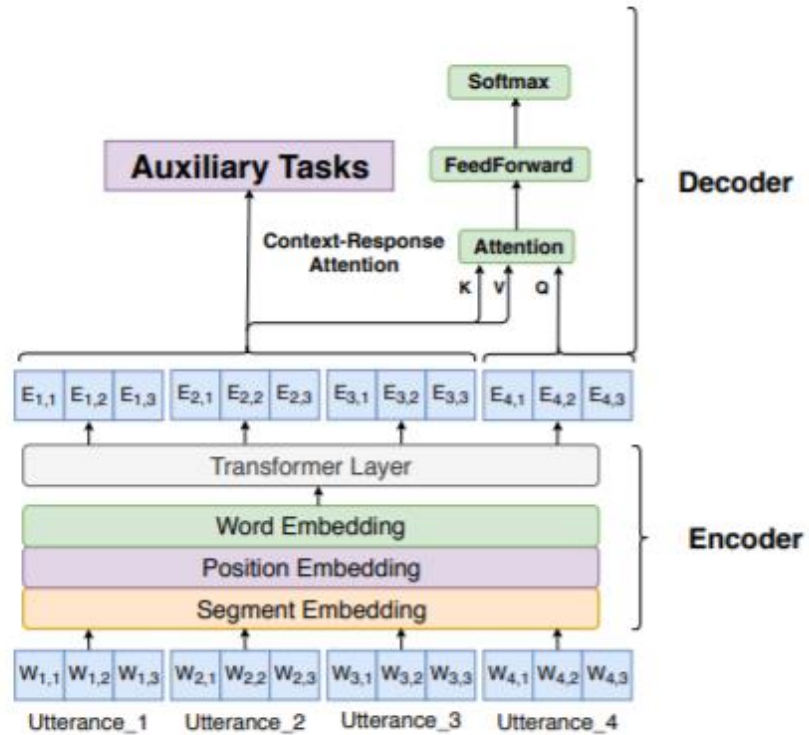


Figure 1: Architecture of the generation model.



# Auxiliary Tasks

- Word order recovery:

**Word order recovery:** Figure 2 (a) illustrates the task. Given a randomly sampled utterance  $U = (w_1, \dots, w_k)$  from a context  $\mathcal{U}$ , we randomly shuffle the words in  $U$  and obtain a disordered utterance  $\bar{U} = (\bar{w}_1, \dots, \bar{w}_k)$ . Then, we replace  $U$  in  $\mathcal{U}$  with  $\bar{U}$  and form a corrupt context  $\bar{\mathcal{U}}$ . The goal of the task is to predict  $U$  from  $\bar{U}$ . The loss of the task can be formulated as

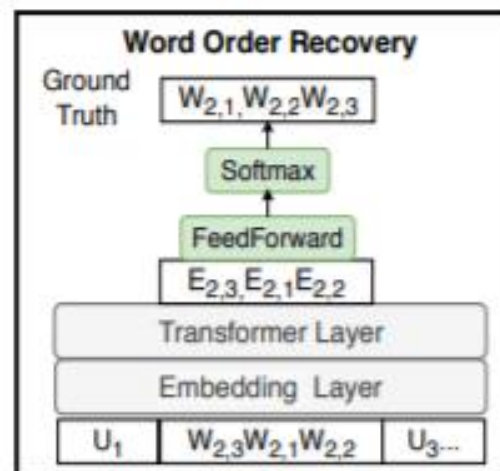
$$\mathcal{L}_{\text{wor}} = -\frac{1}{k} \sum_{i=1}^k \log(p(w_i|\bar{U})), \quad (7)$$

$$p(w_i|\bar{U}) = \text{softmax}(W_s E(\bar{w}_i)),$$

where  $E(\bar{w}_i)$  is obtained from  $E(\bar{\mathcal{U}})$  which is the representation of  $\bar{\mathcal{U}}$  given by the encoder of the generation model,  $W_s$  is shared with Equation (6).

For this task, the mask matrix  $M$  in Equation (4) is defined by:

$$M_{ij} = \begin{cases} 0, & w_i \text{ and } w_j \text{ are in the same utterance,} \\ -\infty, & w_i \text{ and } w_j \text{ are in different utterances.} \end{cases} \quad (8)$$



# Auxiliary Tasks

- Utterance order recovery:

**Utterance order recovery:** Figure 2 (d) illustrates the task. Given context  $\mathcal{U} = (U_1, \dots, U_n)$ , we randomly shuffle the utterances and obtain a disordered context  $\bar{\mathcal{U}} = (U_{o_1}, \dots, U_{o_n})$ . The goal is to predict the correct positions for utterances in  $\bar{\mathcal{U}}$ . The prediction model falls in a read-process-write framework (Vinyals et al., 2015). In the reading module, the model first represents  $\bar{\mathcal{U}}$  as  $\bar{E} = (\bar{E}(w_{1,1}), \dots, \bar{E}(w_{n,m}))$  via the encoder of the generation model, where  $w_{i,j}$  is the  $j$ -th word in utterance  $U_{o_i}$  (words within an utterance are ordered), and then obtains the representation of utterance  $U_{o_i}$  through

$$S_i = \sum_{j=1}^{k_i} \bar{E}(w_{i,j}), \quad (9)$$

where  $k_i$  is the number of words in  $U_{o_i}$ .  $\mathcal{S} = \{S_i\}_{i=1}^n$  forms a sentence memory that is accessible by the processing module. The processing module exploits multi-head self-attention and GRU to guarantee the property that vectors retrieved from memory  $\mathcal{S}$  will not change if the memory is randomly shuffled. Formally, the processing module

We then apply another GRU decoder to decode  $\{o_1, o_2, \dots, o_n\}$  one by one.

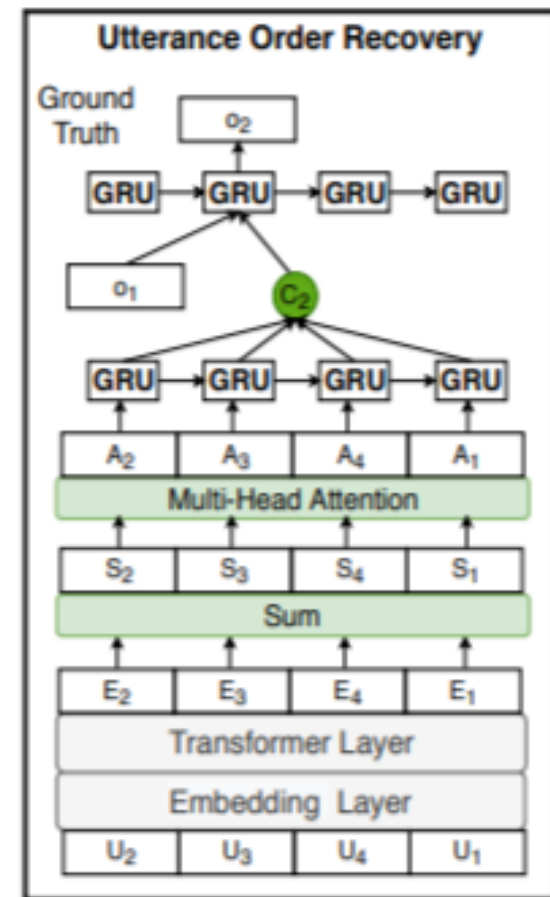
$$\bar{h}_i = \text{GRU}(\bar{h}_{i-1}, [c_i \oplus x_i]), \quad (11)$$

where  $\bar{h}_{i-1}$  is the hidden state at step  $i - 1$  with  $\bar{h}_0 = h_n$ ,  $x_i$  is the embedding of  $o_{i-1}$  (i.e., the embedding of the ground-truth position of  $U_{o_{i-1}}$  in  $\mathcal{U}$ ), and  $c_i$  is a context vector which is defined via attention over  $\{h_t\}_{t=1}^n$ :

$$\begin{aligned} c_i &= \sum_{t=1}^n a_{i,t} h_t, \\ \{a_{i,t}\}_{t=1}^n &= \text{softmax}(\{e_{i,t}\}_{t=1}^n), \\ e_{i,t} &= V^\top \tanh(W_1 \bar{h}_{i-1} + W_2 h_t + b_1), \end{aligned} \quad (12)$$

where  $V_1$ ,  $W_1$ ,  $W_2$ , and  $b_1$  are parameters. The prediction model is finally formulated as

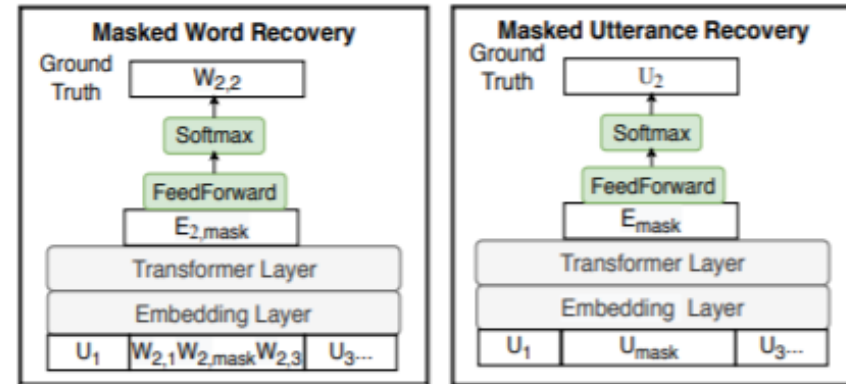
$$\begin{aligned} P(o_i | \{o_1, \dots, o_{i-1}\}, \bar{\mathcal{U}}) &= \text{softmax}(u_i), \\ u_i &= \text{FNN}(\bar{h}_i \oplus x_i \oplus c_i). \end{aligned} \quad (13)$$



# Auxiliary Tasks

- Masked content recovery: word level and Utterance level, Given a context  $U = (w_1, w_2, \dots, w_m)$ , suppose the masked context is  $\bar{U} = (w_1^*, w_2^*, \dots, w_m^*)$ . Where  $w_i^* = [MASK]$  if  $w_i$  is masked, otherwise  $w_i^* = w_i$ . The loss is formulated as:

$$\begin{aligned}\mathcal{L}_x &= -\frac{1}{k} \sum_{i=1}^m \mathbb{I}[w_i^*=[MASK]] \log(p(w_i|\bar{U})), \\ k &= \sum_{i=1}^m \mathbb{I}[w_i^*=[MASK]], \\ p(w_i|\bar{U}) &= \text{softmax}(W_s E(w_i^*)),\end{aligned}\tag{15}$$



# Experiments

- We conduct experiments on DailyDialog (Li et al., 2017), PERSONA-CHAT (Zhang et al., 2018), and the Ubuntu Dialogue Corpus (UDC) (Lowe et al., 2015), and compare our model with state-of-the-art baselines in terms of response quality, parameter size, and decoding speed

# Experiments

Dataset	Model	PPL	BLEU	Distinct-1	Distinct-2	Average	Greedy	Extrema	Parameter size	Decoding speed
DailyDialog	HRED	56.22	0.535	1.553	3.569	81.393	65.546	48.109	34.5M	14.79ms
	HRAN	47.23	0.447	1.953	7.400	83.460	67.239	<b>49.599</b>	38.2M	17.15ms
	VHRED	44.79	0.997	1.299	6.113	83.866	67.186	48.570	34.8M	15.67ms
	SSN	44.28	1.250	2.309	7.266	72.796	<b>73.069</b>	44.260	20.0M	12.69ms
	ReCoSa	42.34	1.121	1.987	10.180	84.763	67.557	48.957	73.8M	40.89ms
	Our Model	<b>38.60</b>	<b>1.658</b>	<b>3.457</b>	<b>14.954</b>	<b>85.224</b>	69.518	49.069	20.3M/14.4M	12.15ms
PERSON-CHAT	HRED	46.04	1.279	0.164	0.450	83.329	64.486	47.132	28.3M	13.14ms
	HRAN	41.94	1.997	0.235	0.771	82.850	65.556	47.882	33.1M	18.43ms
	VHRED	42.07	2.181	0.312	1.915	82.995	65.578	46.810	28.8M	20.27ms
	SSN	47.90	2.288	0.637	2.623	<b>85.002</b>	66.752	47.461	15.2M	15.82ms
	ReCoSa	34.19	2.258	0.915	4.217	83.963	66.498	48.163	68.7M	39.38ms
	Our Model	<b>33.23</b>	<b>2.434</b>	<b>1.279</b>	<b>5.816</b>	83.632	<b>66.778</b>	<b>48.552</b>	18.4M/12.5M	13.89ms
Ubuntu	HRED	58.23	0.874	0.602	2.724	76.187	62.869	37.508	24.1M	25.09ms
	HRAN	48.14	0.922	0.472	2.217	76.654	62.145	37.282	29.5M	31.07ms
	VHRED	52.34	0.906	0.571	2.933	76.496	63.051	36.039	24.7M	30.47ms
	SSN	57.82	<b>1.681</b>	0.557	2.370	76.431	61.597	35.976	12.3M	21.11ms
	ReCoSa	43.67	0.911	0.722	4.439	77.619	<b>63.239</b>	36.742	60.6M	45.34ms
	Our Model	<b>40.94</b>	1.625	<b>0.783</b>	<b>5.151</b>	<b>78.754</b>	62.738	<b>38.538</b>	14.4M/8.5M	22.98ms

Table 2: Evaluation results on automatic metrics. Numbers in bold indicate the best performing model on the corresponding metrics.

