

Unsupervised Knowledge Selection for Dialogue Generation

Xiuyi Chen, Feilong Chen, Fandong Meng, Peng Li, Jie Zhou

Pattern Recognition Center, WeChat AI, Tencent Inc, Beijing, China

{hugheren.chan, ivess.chan}@gmail.com

{fandongmeng, patrickpli, withtomzhou}@tencent.com

Abstract

Knowledge selection is an important and challenging task which could provide the appropriate knowledge for informative dialogue generation. However, the needed gold knowledge label is difficult to collect in reality. In this paper, we study knowledge selection for dialogue generation in the unsupervised scenario and propose a novel Distilled Distant Supervision Loss (DDSL) to supervise knowledge selection when the gold knowledge label is unknown. Specifically, we first obtain an oracle knowledge label via distant supervision and then leverage knowledge distillation to alleviate the noisy labeling problem of distant supervision. Furthermore, we propose a pretraining-finetuning strategy to deal with the mismatch knowledge selection problem that models tend to select the mismatched knowledge for dialogue generation in the unsupervised setting and will cause the degeneration of knowledge-aware decoder. Experiments on two knowledge-grounded dialogue datasets show that our approach manages to select knowledge more accurately in the unsupervised setting and generates more informative responses, even outperforming many strong supervised baselines.¹

1 Introduction

To avoid general and dull dialogue generation (Li et al., 2016), knowledge-grounded dialogue which equips dialogue systems with external knowledge has become a popular research topic. Thanks to the hand-collected knowledge-grounded dialogue datasets which align each dialogue (even each utterance) with a pre-identified document (Zhang et al., 2018; Zhou et al., 2018; Moghe et al., 2018; Zhou et al., 2020), many researchers focus on injecting the given knowledge to generate informative responses and achieve promising results (Yavuz

et al., 2019; Tang and Hu, 2019; Qin et al., 2019a; Li et al., 2019b; Zheng and Zhou, 2019; Meng et al., 2019; Ren et al., 2019; Ye et al., 2020; Lin et al., 2020). However, they usually need the pre-identified knowledge and the knowledge access task is less studied (Kim et al., 2020b) which is the precursor to knowledge dialogue generation in reality (Zheng et al., 2020).

It is natural to extract the external knowledge via information retrieval technology. Several works regard the retrieved knowledge sentences as the pre-identified document (Ghazvininejad et al., 2018; Michel Galley, 2018; Yang et al., 2019b; Gopalakrishnan et al., 2019). However, the retrieved document contains redundant and irrelevant information which are harmful for dialogue generation (Zhao et al., 2020b). Hence, knowledge selection which chooses an appropriate sentence from the pre-retrieved knowledge pool gains much attention and it plays the role of knowledge access for generation. Dinan et al. (2019) first propose knowledge selection for dialogue generation which are two sequential subtasks and the generation is based on the selected knowledge. Several works follow their setting and achieve improvements with latent variable models (Kim et al., 2020a; Chen et al., 2020b) or more complex selection mechanism (Niu et al., 2019; Meng et al., 2020; Zheng et al., 2020; Meng et al., 2021). Although those works show promising performance with explicit use of knowledge in open-domain dialogue, they still need gold knowledge labels to train the selection module well (Kim et al., 2020a). And it is still less studied to make knowledge selection work well without gold knowledge label, which is valuable and challenging.

In this paper, we explore knowledge selection for dialogue generation in the unsupervised scenario and propose a novel Distilled Distant Supervision Loss (DDSL) to supervise knowledge selection when there is no gold knowledge label. Specifically, we first obtain an oracle knowledge

¹The codes and model checkpoints will be available at <https://github.com/ErenChan/UKSDG>

label via distant supervision (Mintz et al., 2009) to substitute the gold one in the unsupervised setting. However, distant supervision inevitably suffers from the noisy labeling problem due to literally matching (Yang et al., 2019a). **Therefore, to train knowledge selection well without gold label, we leverage knowledge distillation to reduce the noise of the oracle label.** Furthermore, we find that models tend to select the mismatched knowledge for dialogue generation in the unsupervised setting. And forcing the knowledge-aware decoder to leverage the selected knowledge at training will cause the decoder degenerating into the knowledge-independent decoder. To deal with this problem, **we propose to pretrain the knowledge selection and response generation independently and then finetune the decoder with the selected knowledge using different sample weighting scores.** We demonstrate the effectiveness of our unsupervised approach on two knowledge-grounded dialogue datasets, i.e., Wizard of Wikipedia (Dinan et al., 2019) and Holle (Moghe et al., 2018) in comparison with various supervised and unsupervised baselines.

Our contributions are summarized as follows:

- We propose Distilled Distant Supervision Loss to make knowledge selection work well in the unsupervised scenario where the gold knowledge label is not available.
- We propose a pretraining-finetuning strategy to alleviate the degeneration of knowledge-aware decoder caused by the mismatch knowledge selection problem.
- Results on two datasets show that our approach manages to select knowledge more accurately in the unsupervised setting and even generates more informative responses than many strong supervised baselines.

2 Approach

2.1 Task Formulation

Given the utterance x_t at each turn t and the associated knowledge pool $\mathcal{K}_t = \{k_t^i\} = \{k_t^1, \dots, k_t^L\}^2$ containing L retrieved candidate sentences k_t^i , the final goal is to generate an informative response y_t . Following (Dinan et al., 2019), we first learn to select the appropriate knowledge k_t^s from the knowledge pool \mathcal{K}_t and then generate the response

² k_t^i, k_t^s, k_t and k_t' indicate **any**, **selected**, **gold** and **oracle** knowledge sentences or labels, respectively.

y_t by incorporating the selected knowledge. In the conventional supervised setting, there exists gold knowledge label to supervise knowledge selection. However, the manually labeled knowledge is difficult to obtain in reality (Lian et al., 2019). As a result, we study the unsupervised knowledge selection for dialogue generation in this paper, which is very valuable and challenging.

2.2 Architecture Overview

In the following subsections, we first introduce the three major components (Section 2.3 ~ 2.5): Encoder, Knowledge Selection (KS) and Decoder, which are trained jointly with **the gold knowledge loss and response generation loss** in the conventional supervised setting, as Figure 1 (a) shows. Then we introduce our Distilled Distant Supervision Loss (DDSL) in Section 2.6 to train knowledge selection well in the unsupervised setting, which consists of distant supervision and knowledge distillation, as Figure 1 (b) shows. Finally, we detail the mismatch knowledge selection problem and how to make the decoder leverage the selected knowledge k_t^s well in Section 2.7.

2.3 Encoder

For each sentence st , we obtain the context aware word representations \mathbf{H}_{st} via BERT (Devlin et al., 2019) and the corresponding sentence representation \mathbf{h}_{st} via Mean Pooling (Cer et al., 2018):

$$\begin{aligned} \mathbf{H}_{st} &= \text{BERT}(st) \in \mathbb{R}^{N_{st} \times d} \\ \mathbf{h}_{st} &= \text{Mean}(\mathbf{H}_{st}) \in \mathbb{R}^d \end{aligned}, \quad (1)$$

where N_{st} is the sentence length and d is the hidden size. Specifically, we represent the utterance x_t with \mathbf{H}_{x_t} and \mathbf{h}_{x_t} , and represent each knowledge sentence $k_t^i \in \mathcal{K}_t$ with $\mathbf{H}_{k_t^i}$ and $\mathbf{h}_{k_t^i}$.

2.4 Knowledge Selection

In this paper, we mainly focus on knowledge selection in the unsupervised setting and adopt the standard dot-product attention over the knowledge candidates to select knowledge (Dinan et al., 2019).

Selection Query: The selection query consists of the current utterance, the dialogue history $dh_t = [x_1, y_1, \dots, x_{t-1}, y_{t-1}]$ and the history of selected knowledge $kh_t = [k_1', \dots, k_{t-1}']$ as they help the knowledge selection (Kim et al., 2020a). Formally, we use two GRUs (Cho et al., 2014) to summarize the dialogue and knowledge selection history as the

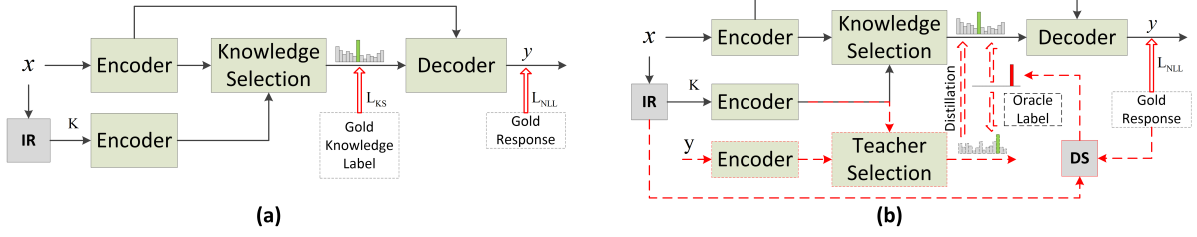


Figure 1: The framework of knowledge selection for dialogue generation. (a) In the supervised setting. (b) In the unsupervised setting. Specifically, we replace the gold knowledge loss with our Distilled Distant Supervision Loss (DDSL) which contains Distant Supervision (DS) and knowledge distillation (see the dotted red lines).

corresponding state vectors s_{dh_t} and s_{kh_t} :

$$\begin{aligned} s_{dh_t} &= \text{GRU}_{dh}([\mathbf{h}_{x_t}; \mathbf{h}_{y_t}], s_{dh_{t-1}}) \in \mathbb{R}^d \\ s_{kh_t} &= \text{GRU}_{kh}(\mathbf{h}_{k'_t}, s_{kh_{t-1}}) \in \mathbb{R}^d \end{aligned} \quad (2)$$

where s_{dh_0} and s_{kh_0} are zero vectors, \mathbf{h}_{x_t} , \mathbf{h}_{y_t} and $\mathbf{h}_{k'_t}$ are sentence vectors of utterance x_t , response y_t and the oracle knowledge k'_t (will be described in Equation 8) and $[\cdot; \cdot]$ denotes concatenation. Then, the selection query \mathbf{q}_t is obtained:

$$\mathbf{q}_t = \mathbf{W}_q [s_{kh_{t-1}}; s_{dh_{t-1}}; \mathbf{h}_{x_t}] \in \mathbb{R}^d. \quad (3)$$

Knowledge Selection: The knowledge selection distribution $\mathcal{S} \in \mathbb{R}^L$ over the knowledge pool $\mathcal{K}_t \in \mathbb{R}^L$ is obtained by the dot-product attention:

$$\mathcal{S}_{k_t^i} = \frac{\exp(\mathbf{h}_{k_t^i} \cdot \mathbf{q}_t)}{\sum_{k_t^j \in \mathcal{K}_t} \exp(\mathbf{h}_{k_t^j} \cdot \mathbf{q}_t)}. \quad (4)$$

Finally, the knowledge k_t^s with the highest attention score is selected for further response generation. If the gold knowledge k_t exists, we could train this task via the Cross Entropy (CE) loss:

$$\mathcal{L}_{KS} = \mathcal{L}_{CE}(\mathcal{S}, k_t) = -\log \mathcal{S}_{k_t}, \quad (5)$$

2.5 Decoder

Following (Dinan et al., 2019; Kim et al., 2020a), our Transformer-based decoder takes the representation concatenation $\mathbf{H}_{rc} = [\mathbf{H}_{x_t}; \mathbf{H}_{k_t^s}] \in \mathbb{R}^{N_{rc}, d}$ of current utterance x_t and the selected knowledge k_t^s as input, and uses the copy mechanism (Gu et al., 2016) to generate responses. The process of generating a word can be formulated as follows:

$$\begin{aligned} \mathbf{s}_t^n &= \text{TD}(\mathbf{H}_{rc}, y_t^{<n}) \in \mathbb{R}^d \\ \mathbf{p}_v &= \text{softmax}(\mathbf{W}_o \mathbf{s}_t^n) \in \mathbb{R}^{|\mathcal{V}|} \\ \mathbf{p}_{cp}, \tilde{\mathbf{s}}_t^n &= \text{MultiHead}_{cp}(\mathbf{s}_t^n, \mathbf{H}_{rc}, \mathbf{H}_{rc}) \quad , \quad (6) \\ p_{gen} &= \text{sigmoid}(\mathbf{W}_{gen} \tilde{\mathbf{s}}_t^n) \in \mathbb{R}^1 \\ \mathbf{p}(\mathcal{V}) &= p_{gen} \cdot \mathbf{p}_v + (1 - p_{gen}) \cdot \mathbf{p}_{cp} \in \mathbb{R}^{|\mathcal{V}|} \end{aligned}$$

where TD denotes the Transformer decoder, \mathbf{s}_t^n is the hidden vector for the n -th word in the response y_t at t -th turn, \mathbf{p}_{cp} is the attention weight of the first head in the additional multi-head self-attention layer for the copy mechanism, which is short for MultiHead_{cp} (Vaswani et al., 2017), \mathcal{V} is the vocabulary, and $\mathbf{p}(\mathcal{V})$ is the final generation distribution. Finally, we generate the word y_t^n with the highest probability, and we keep generating by feeding y_t^n to the decoder until the “<eos>” token is generated. We train the generation task the Negative Log-Likelihood (NLL) loss:

$$\mathcal{L}_G = \mathcal{L}_{NLL} = -\log p(y_t | x_t, k_t^s) \quad (7)$$

The model is trained with the loss $\mathcal{L} = \mathcal{L}_{KS} + \mathcal{L}_G$, where \mathcal{L}_{KS} is the knowledge selection loss, i.e., Equation 5 in the supervised setting.

2.6 Distilled Distant Supervision Loss

In this section, we will introduce our Distilled Distant Supervision Loss (DDSL) to train knowledge selection well in the unsupervised setting, which consists of distant supervision, label weighting and knowledge distillation. **Actually, we first obtain a noisy oracle knowledge label via distant supervision.** And then our DDSL tries to reduce the noise via label weighting and knowledge distillation.

Distant Supervision: Suppose that we have the utterance x_t , response y_t and the retrieved knowledge pool \mathcal{K}_t without knowing the gold knowledge label, we first calculate the confidence score $\mathcal{W}_{k_t^i}$ whether each knowledge k_t^i is matched up with this dialogue flow by:

$$\begin{aligned} \text{F1}(a, b) &= \frac{2 \cdot |\text{set}(a) \cap \text{set}(b)|}{|\text{set}(a)| + |\text{set}(b)| + \epsilon}, \quad (8) \\ \mathcal{W}_{k_t^i} &= \text{softmax}_{\tau}(\text{F1}(k_t^i, y_t)) \end{aligned}$$

where $\text{set}(a)$ or $\text{set}(b)$ indicates the tokens in the string a or b , $\text{softmax}_{\tau}(z_i) = e^{(z_i/\tau)} / \sum_j e^{(z_j/\tau)}$

and τ is the temperature to reshape the confidence probability distribution $\mathcal{W} \in \mathbb{R}^L$ over the knowledge pool. Then, we obtain the oracle knowledge k'_t with the highest confidence score, **assuming that the gold knowledge should contribute most tokens to the response generation**. This is possible because that 1) the causal modeling specified conditions (Selltiz et al., 1976) hold between knowledge selection and response generation (Tuan et al., 2020), and 2) it is common for humans to (involuntarily) produce utterances which are copied or suitably modified from background knowledge sentences (Moghe et al., 2018).

Although we can directly replace the gold label k_t in Equation 5 with the alternative one k'_t , there are some noise. Therefore, we modify Equation 5 with label weighting via the confidence score:

$$\mathcal{L}_{KS} = \mathcal{L}_{WCE}(\mathcal{S}) = \mathcal{L}_{CE}(\mathcal{S}, k'_t) \cdot \mathcal{W}_{k'_t}. \quad (9)$$

Knowledge Distillation: We further alleviate the noisy labeling problem of distance supervision via Knowledge Distillation (KD) as shown in Figure 1 (b). Following (Tian et al., 2020; Chen et al., 2020b), **the teacher takes the context and response as input and generates the distribution of knowledge selection as soft target**. Compared with the student, i.e., the standard knowledge selection module described in Section 2.4, **teacher has the gold response as an additional input**.

Specifically, we make up the teacher query as $\mathbf{q}_t^{\text{tea}} = \mathbf{W}_{\text{tea}}[\mathbf{s}_{dh_t}; \mathbf{s}_{kh_{t-1}}] \in \mathbb{R}^d$, which contains more information (i.e., the response) than \mathbf{q}_t in Equation 3. Then we use this query $\mathbf{q}_t^{\text{tea}}$ to obtain the teacher’s knowledge selection distribution $\mathcal{T} \in \mathbb{R}^L$ by Equation 4. Finally, online response-based knowledge distillation³ is formulated as:

$$\begin{aligned} \mathcal{L}_{KD} &= \mathcal{L}_{WCE}(\mathcal{T}) + D_{KL}(\mathcal{T} \parallel \mathcal{S}) \\ &= \mathcal{L}_{CE}(\mathcal{T}, k'_t) \cdot \mathcal{W}_{k'_t} + \sum_{k_t^i \in \mathcal{K}_t} \mathcal{T}_{k_t^i} \log \frac{\mathcal{T}_{k_t^i}}{\mathcal{S}_{k_t^i}}, \end{aligned} \quad (10)$$

where the first term is used for teacher training and the second term transfers the knowledge from teacher to student based on the Kullback-Leibler (KL) divergence between the teacher and student distributions (i.e., $\mathcal{T} \in \mathbb{R}^L$ and $\mathcal{S} \in \mathbb{R}^L$).

Although the teacher is also trained with the noisy label, the teacher produces an independent

³As surveyed in (Gou et al., 2020), response-based knowledge usually refers to the neural response of the last output layer of the teacher model, and online distillation means we train the teacher and student together.

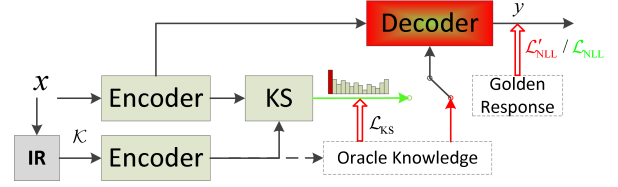


Figure 2: To deal with the mismatch problem, we first pretrain the Knowledge Selection (KS) and the decoder with the matched knowledge⁴ in parallel along red lines. Then we finetune the decoder with the selected knowledge weightedly along the green line.

source of variance that can be used to cancel out the variance introduced by the label noise (Li et al., 2017). Moreover, previous works have proved **that the student can still be enhanced by the noisy teacher (Sau and Balasubramanian, 2016; Xie et al., 2020; Yuan et al., 2020)**. Therefore, we believe that the student trained by two supervision signals gets benefits from the regularization of the soft target (Yuan et al., 2020).

To sum up, our DDSL loss is as follows:

$$\mathcal{L}_{KS} = \mathcal{L}_{DDSL} = \mathcal{L}_{WCE}(\mathcal{S}) + \mathcal{L}_{KD}, \quad (11)$$

which consists of distant supervision, label weighting in Equation 9 and knowledge distillation in Equation 10 for unsupervised knowledge selection.

2.7 Training

The Mismatch Knowledge Selection Problem:

As we know, there are chances that the selected knowledge is not the gold one due to 1) the diversity of knowledge selection in conversation (Kim et al., 2020a) and 2) the under-optimized knowledge selection at early training stage (Zhao et al., 2020b). And it is more serious in the unsupervised setting where it is hard to train knowledge selection well. The mismatch knowledge selection problem occurs due to the training paradigm in Figure 1 where the decoder is trained to generate the gold response with mismatched knowledge. **This mismatch problem causes the knowledge-aware decoder to take the selected knowledge as noise and degenerate into the knowledge-independent decoder.**

Our Pretraining-Finetuning Strategy: Although training the decoder with the matched knowledge⁴ solves the mismatch problem. It can’t deal with wrong knowledge selection at inference which is often the case, yet never seen at training.

⁴Note that the oracle knowledge is most literally matched with the gold response by the metric defined in Equation 8.

We take the plain idea as our pretraining stage, and then train the decoder to adapt to the selected knowledge using different weighting scores in the finetuning stage.

In the pretraining stage, we train knowledge selection and response generation in parallel as Figure 2 shows. The pretraining loss is as follows:

$$\begin{aligned}\mathcal{L}'_{\text{NLL}} &= -\log p(y_t|x_t, k'_t) \\ \mathcal{L} &= \mathcal{L}_{\text{KS}} + \mathcal{L}'_{\text{NLL}}\end{aligned}\quad (12)$$

where we use $\mathcal{L}_{\text{KS}} = \mathcal{L}_{\text{DDSL}}$ in the unsupervised setting and the decoder is trained to generate the gold response with the oracle knowledge k'_t instead of the selected one k_t^s . Therefore, the decoder learns how to incorporate the matched knowledge⁴ k'_t into the response generation because k'_t is much more accurate than the selected one k_t^s as Table 4 shows. And we could alleviate the mismatch problem from the pretraining process because 1) we get a fully optimized knowledge selection module and 2) the pretrained decoder provides a good initialization for finetuning.

In the finetuning stage, we continually train the pretrained decoder to adapt to the pretrained knowledge selection module with the sample weighting idea. And the finetuning loss is defined as follows:

$$\mathcal{L}_G = \mathcal{L}_{\text{NLL}} \cdot (1 + \mathcal{W}_{k_t^s}) \quad (13)$$

where \mathcal{L}_{NLL} is defined in Equation 7 and $\mathcal{W}_{k_t^s}$ is the confidence or weighting score of the selected knowledge k_t^s defined in Equation 8. As mentioned above, the mismatch knowledge selection problem is caused by the training paradigm that **we may train the decoder to generate the gold response with the mismatched knowledge from the knowledge selection module**. Here, we finetune the pretrained decoder with the selected knowledge k_t^s by giving higher importance weights if the selected knowledge is suitable for the gold response generation. In this way, we further alleviate the mismatch problem because we highlight the matched samples by assigning an importance weight to each instance (x_t, k_t^s, y_t) to reform the training data (Cai et al., 2020; Dong et al., 2020).

3 Experiments

3.1 Dataset

We evaluate our model on two public knowledge-grounded dialogue datasets: Wizard of Wikipedia (WoW) (Dinan et al., 2019) and Holl-E (Moghe

et al., 2018), both of which provide the knowledge candidates with gold knowledge labels for knowledge selection. To test our approach in the unsupervised setting, we do not use the gold knowledge label provided in those datasets.

WoW contains the dialogues between two participants on some open-domain topics, where one is a curious learner while the other plays the role of a knowledgeable expert with access to the knowledge pool. Each knowledge pool contains about 61 sentences on average per turn, which are retrieved from Wikipedia based on the dialogue context via the IR system. There are 18430, 1948 and 1933 dialogues for training, validation and test, respectively. According to the topic overlapping, the test set is split into two subsets: 965 Test Seen and 968 Test Unseen dialogues, where Test Unseen consists of 58 topics never seen in train or validation.

Holl-E contains 7228, 930 and 913 dialogues for training, validation and test, respectively. Two test versions are provided: one with a single gold reference, the other with multiple gold references (more than one gold knowledge sentences and corresponding responses for each given conversation context). Each dialogue is assigned with a document of about 60 sentences on average as the knowledge pool. Here, we use the modified version (Kim et al., 2020a) which fits for knowledge selection.

3.2 Models for Comparison

We compare our method with a set of baselines:⁵

3.2.1 No Knowledge

S2S_{Transformer} is a Seq2Seq model based on Transformer (Vaswani et al., 2017) that does not leverage the external knowledge.

S2S_{BERT} replaces the Transformer Encoder with a pretrained BERT (Devlin et al., 2019).

3.2.2 Supervised Knowledge Selection

TMN is short for End-to-End Transformer MemNet (Dinan et al., 2019), which selects knowledge based on the Transformer memory network and generate responses via the Transformer decoder.

TMN_{BERT+PostKS+CP}, implemented by (Kim et al., 2020a), enhances the encoder with BERT, knowledge selection module with PostKS (Lian et al., 2019) and decoder with the copy mechanism (CP).

⁵See Appendix A.1 for more baselines, from knowledge-aware generation (Huang et al., 2020) to knowledge selection in the supervised, semi-supervised and unsupervised settings.

Setting	Row	Method	Test Seen				Test Unseen			
			Acc \uparrow	PPL \downarrow	R-1 \uparrow	R-2 \uparrow	Acc \uparrow	PPL \downarrow	R-1 \uparrow	R-2 \uparrow
0 No Knowledge	0_0	S2S $^{\dagger}_{\text{Transformer}}$ (Vaswani et al., 2017)	N/A	57.4	16.5	4.4	N/A	102.0	14.6	2.9
	0_1	S2S $^{\dagger}_{\text{BERT}}$	N/A	53.9	17.2	4.8	N/A	93.3	14.9	3.3
1 Supervised	1_0	TMN (Dinan et al., 2019)	21.1	63.5	16.9	N/A	14.3	97.3	14.4	N/A
	1_1	TMN $_{\text{BERT+PostKS+CP}}$ (2020a)	25.5	52.2	19.0	6.5	14.4	83.4	15.6	3.9
	1_2	SKT (Kim et al., 2020a)	26.8	52.0	19.3	6.8	18.3	81.4	16.1	4.2
2 Unsupervised	2_0	TMN $_0$ (Dinan et al., 2019)	13.4	66.5	15.9	N/A	11.8	103.6	14.3	N/A
	2_1	PostKS (Lian et al., 2019)	4.8	79.1	13.0	1.0	4.2	193.8	13.1	1.0
	2_2	SKT $_0$ (Kim et al., 2020a)	0.3	54.7	17.1	4.6	0.1	88.2	15.5	3.4
	0	UKSDG w/o DDSL	5.2	49.3	17.4	5.1	5.1	82.8	15.1	3.3
	1	UKSDG $_{\text{vec}}$ w/o DDSL	13.5	46.1	17.8	5.2	13.1	82.3	15.5	3.3
	2	UKSDG (ours)	23.8	51.8	19.5	6.8	16.2	76.3	16.3	4.4
	3	UKSDG $_{\text{PF}}$ w/o SW (ours)	26.4	44.1	20.3	7.4	20.8	64.5	17.8	5.6
	4	UKSDG $_{\text{PF}}$ (ours)	26.4	45.0	20.6	7.7	20.8	65.5	18.2	5.9

Table 1: Quantitative results on WoW. Our approach manages to select knowledge more accurately in the unsupervised setting and generate more informative responses than the strong baselines in the supervised setting. Note that models with “ \dagger ” are implemented by ourselves and other models with citation are from the original paper.

SKT (Kim et al., 2020a), short for Sequential Knowledge Transformer, uses the posterior distribution by sequential latent modeling and achieves promising results in the supervised setting.

3.2.3 Unsupervised Knowledge Selection

TMN $_0$ is TMN, trained only via generation loss in Equation 7 without knowledge loss in Equation 5. **SKT $_0$** is SKT optimized without knowledge loss. **PostKS** (Lian et al., 2019) takes the benefits of latent variable models and leverages the posterior knowledge distribution as a pseudo label for knowledge selection without knowledge loss. Here, we use the results provided by (Kim et al., 2020a).

3.2.4 Our Unsupervised Methods

We implement our model in the unsupervised setting, namely Unsupervised Knowledge Selection for Dialogue Generation (**UKSDG**), which is optimized with our DDSL in Equation 11 for unsupervised knowledge selection and generation loss in Equation 7. **UKSDG $_{\text{PF}}$** indicates that we adopt our Pretraining-Finetuning strategy to alleviate the mismatch knowledge selection problem in Section 2.7. Furthermore, we remove several components for ablation study:

- (1) **UKSDG w/o DDSL** is only optimized by the generation loss in Equation 7 without our DDSL.
- (2) **UKSDG $_{\text{vec}}$ w/o DDSL** further replaces the decoder input \mathbf{H}_{rc} (in Section 2.5) with the averaged knowledge vector enhanced context $\mathbf{H}_{vec} = \mathbf{H}_{x_t} + \sum_{k_t^i \in \mathcal{K}_t} \mathcal{S}_{k_t^i} \cdot \mathbf{h}_{k_t^i}$.
- (3) **UKSDG $_{\text{PF}}$ w/o SW** does not use the Sample Weighting in Equation 13.

3.3 Implementation Details

We use TensorFlow 2.0 to implement our approach base on SKT⁶. All sentences are encoded by the shared BERT $_{\text{BASE}}$ (Devlin et al., 2019), and the response is greedily generated via a 5-layer Transformer decoder with copy mechanism. The hidden size d is 768 and the vocabulary size $|V|$ is 30,522. The knowledge selection module contains two separate one-layer GRUs and one projection layer. Our proposed DDSL contains no trainable parameters except one projection layer in the teacher selection module. And the temperature τ is 0.1.

We use the Adam optimizer (Kingma and Ba, 2014) with gradient clipping at 0.4 to train our models on a single GPU (TITAN Xp). The learning rate is $2e^{-5}$ and the batch⁷ size is 1. Moreover, we apply label smoothing (Pereyra et al., 2017) and set 0.1 for knowledge selection and 0.05 for response generation. In the pretraining-finetuning strategy, we use 5 and 20 epochs in the pretraining and finetuning stage, respectively. The pretrained models are selected according to the accuracy score and other models are selected according to the R-1 score since knowledge selection aims to serve for high-quality generation.

It takes almost the same time for the convergence of UKSDG and KSDG as we only replace our DDSL with the gold knowledge selection loss. The convergence of SKT is a bit slower as it is hard

⁶Thanks for their processed datasets, models and evaluation codes at <https://github.com/bckim92/sequential-knowledge-transformer>.

⁷Each example is a dialogue rather than an individual turn.

Setting	Row	Method	Single Reference				Multi Reference			
			Acc \uparrow	PPL \downarrow	R-1 \uparrow	R-2 \uparrow	Acc \uparrow	PPL \downarrow	R-1 \uparrow	R-2 \uparrow
0 No Knowledge	0_0	S2S [†] _{Transformer} (Vaswani et al., 2017)	N/A	105.8	19.1	9.8	N/A	72.4	24.1	12.9
	0_1	S2S [†] _{BERT}	N/A	89.4	19.0	9.6	N/A	61.9	23.5	12.1
1 Supervised	1_0	TMN (Dinan et al., 2019)	22.7	140.6	20.1	10.3	32.3	83.6	24.3	12.8
	1_1	TMN _{BERT+PostKS+CP} (2020a)	27.8	47.4	29.2	22.3	37.8	27.9	35.9	29.0
	1_2	SKT (Kim et al., 2020a)	29.2	48.9	29.8	23.1	39.2	28.5	36.5	29.7
2 Unsupervised	2_1	PostKS (Lian et al., 2019)	1.5	196.6	15.2	6.0	3.2	114.1	19.2	7.9
	2_2	SKT ₀ [‡] (Kim et al., 2020a)	0.1	77.1	19.2	9.9	0.1	75.8	19.4	10.1
	0	UKSDG w/o DDSL	2.9	76.5	20.0	10.5	4.3	52.0	25.0	13.6
	1	UKSDG _{vec} w/o DDSL	19.1	78.1	20.9	10.9	28.8	53.2	25.4	13.3
	2	UKSDG (ours)	24.2	55.4	29.7	23.1	33.5	31.9	36.4	29.7
	3	UKSDGPF w/o SW (ours)	25.1	38.7	30.8	23.2	35.0	22.7	37.9	30.0
	4	UKSDG _{PF} (ours)	25.1	39.4	31.0	24.0	35.0	22.9	38.1	31.1

Table 2: Quantitative results on Holl-E. The method with “[‡]” is reported by rerunning the released code, and models with “[†]” are implemented by ourselves.

to optimize the latent variable model. It takes about 2.5 times as long for the convergence of UKSDG_{PF} and KSDG_{PF}. The computation for confidence calculation during inference is the same for UKSDG, KSDG (w/ and w/o PF) and SKT because we adopt the same backbone. All of our models contain about 174M parameters.

3.4 Evaluation

Automatic Evaluation. We automatically evaluate knowledge selection with accuracy (Acc), response generation with perplexity (PPL), unigram F1 (R-1) and bigram F1 (R-2), which are commonly used in this task (Dinan et al., 2019; Kim et al., 2020a; Chen et al., 2020b). We also remove all the punctuation and (a, an, the) to compute the R-1 and R-2 scores as (Kim et al., 2020a) do. Note that lower PPL and higher R-1 and R-2 scores indicate better generation quality.

Human Evaluation. We firstly select 100 samples from each test set on WoW for human evaluation. Then we follow (Kim et al., 2020a; Li et al., 2019a) and ask three annotators to evaluate the generation quality according to Engagingness and Knowledgeability from 1 to 4, where 1 means not at all, 2 is a little, 3 is somewhat, and 4 is a lot. Engagingness measures how much do you like the response and Knowledgeability measures the informativeness in the responses.

4 Results and Analysis

4.1 Main Results

Quantitative Results. We report the automatic results on WoW and Holl-E in Table 1 and Table 2,

respectively. And we have the following consistent observations:⁸ (1) Comparing SKT and SKT₀, we see that the gold knowledge loss plays an important role to train knowledge selection well. (2) Comparing row 0 and 2, we see that our proposed DDSL is the key to train knowledge selection well in the unsupervised setting and can be the alternative of the gold knowledge loss. As a result, our approach significantly outperforms the other unsupervised methods on all metrics (significance tests (Koehn, 2004), p-value < 0.01). (3) Although UKSDG_{vec} w/o DDSL could learn some patterns of knowledge selection from the gradient of generation loss, we see that compressing the knowledge into a vector will loss much information for dialogue generation. (4) Although our UKSDG_{PF} usually makes knowledge selection worse than the supervised SKT in row 1_2, we achieve higher generation quality, which indicates that our pretraining-finetuning strategy could alleviate the mismatch knowledge selection problem and emphasizes the importance of leveraging the selected knowledge properly for future study on this task. (5) Comparing row 3 and 4, we see that the sample weighting also helps in the finetuning stage, though PPL score is slightly worse due to the difficulty of injecting the knowledge into responses. (6) Moreover, our approach demonstrates the stronger ability of generalization with smaller performance gap between Test Seen and Test Unseen in Table 1, which we attribute to our DDSL because it works in the unsupervised setting and is more suitable for Test Unseen. For example, compared with SKT in row 1_2, UKSDG_{PF} in row

⁸More observations with other baselines can be found in Appendix A.2.

Method	Test Seen		Test Unseen	
	EGG	KLD	EGG	KLD
SKT	2.70±0.05	2.64±0.05	2.46±0.05	2.50±0.06
UKSDG (ours)	2.49±0.05	2.37±0.06	2.33±0.06	2.27±0.05
UKSDG _{PF} (ours)	2.72±0.05	2.71±0.05	2.49±0.06	2.60±0.05
Gold Response	3.25±0.04	3.37±0.05	3.16±0.05	3.24±0.05

Table 3: Results of human evaluations on WoW. **EGG** and **KLD** denote Engagingness and Knowledgeability, respectively.

4 achieves the highest selection accuracy in Test Unseen with a slightly lower accuracy in Test Seen.

Qualitative Results. We report human evaluation results of the generated responses according to Engagingness and Knowledgeability in Table 3. Comparing UKSDG and UKSDG_{PF}, we see the effectiveness of our pretraining-finetuning strategy which could alleviate the mismatch problem described in Section 2.7. Our UKSDG_{PF} in the unsupervised setting generates responses slightly better than SKT in the supervised setting. Moreover, the improvement is much obvious according to Knowledgeability, which also indicates the importance of using the selected knowledge properly.

4.2 Ablation Study

We have introduced the components of our DDSL in Section 2.6 and shown its effectiveness in Section 4.1. Here, we analyse our DDSL via an ablation study in Table 4. Actually, we train UKSDG_{PF} on knowledge selection task, using our DDSL with components removed. We have the following observations: (1) We see that most of the oracle knowledge from distant supervision is the same as the gold knowledge label by human. Therefore, it is acceptable to directly use this oracle knowledge label to substitute the gold label as the supervision signal (see the last row of Table 4). (2) However, distant supervision inevitably suffers from the noisy labeling problem due to literally matching. For example, there are about 30% oracle knowledges different from the gold ones on WoW where the responses convey knowledge much more flexibly. (3) Loss weighting in Equation 9 helps on WoW where the noisy labeling problem is serious. (4) Knowledge distillation in Equation 10 could further alleviate the label noise, and our method manages to select knowledge accurately in the unsupervised setting.

4.3 Case Study

Table 5 presents several examples of generated responses on WoW and we have the following ob-

Method	WoW		Holl E	
	Seen	Unseen	Single	Multi
Gold label (k_t)	100.0	100.0	100.0	100.0
Oracle Label (k'_t)	70.1	69.0	90.3	90.4
$\mathcal{L}_{\text{DDSL}}$ in Eq. 11	26.4	20.8	25.1	35.0
w/o \mathcal{L}_{KD} in Eq. 10	25.8	19.1	24.1	33.5
w/o \mathcal{L}_{KD} & $\mathcal{W}_{k'_t}$ in Eq. 9	25.5	18.8	24.0	34.2

Table 4: Ablation study. Knowledge selection accuracy of UKSDG_{PF} trained with different losses.

servations: (1) The oracle knowledge from distant supervision contains several informative tokens which could help dialogue generation since they appears in the gold response as defined by Equation 8. In particular, the oracle knowledge in the case 3 is also appropriate and the gold response contains much more information than the gold knowledge, which indicates the diversity of knowledge selection and selecting one sentence may be not enough for informative generation. (2) However, some oracle knowledge labels are different from the gold ones (i.e., the selected one by human). Our model still learns to select knowledge as human, which we attribute to our DDSL since knowledge distillation alleviates the noisy labeling problem of distant supervision. (3) SKT does not leverage the selected knowledge well and generates the dull response. For example, SKT repeats the context in case 1, generates the verbose and contradictory response in case 2 and does not provide new information in case 3. (4) Whereas, our UKSDG_{PF} firstly selects the appropriate knowledge as human does, and then generate fluent and informative responses by alleviate the mismatch knowledge selection problem with the help of the pretraining-finetuning strategy. This indicates the importance of leveraging the selected knowledge properly for future study.

5 Related Work

External knowledge has been wildly explored to enhance dialogue understanding and/or improve dialogue generation (Zhu et al., 2017; Liu et al., 2018; Chen et al., 2019; Wu et al., 2019; Chen et al., 2020a; Tuan et al., 2020; Sun et al., 2020; Zhang et al., 2020; Yu et al., 2020; Ni et al., 2021). To make use of knowledge, knowledge access is very important. Therefore, knowledge selection which selecting appropriate knowledge given the dialog context gains much attention (Zhang et al., 2019; Meng et al., 2019; Ren et al., 2019; Dinan

Case 1	Context		Are you a fan of elvis presley?
	Human	KS	Regarded as one of the most significant cultural icons of the 20th century, he is often referred to as the "king of rock and ..."
		DG	You mean the king of rock and roll . Actually yes I am. best of all time. Don't you agree?
	SKT	KS	Elvis Aaron Presley (January 8, 1935 - August 16, 1977) was an American singer, musician, and actor.
		RG	Yes, I love Elvis Presley!
	UKSDG _{PF}	OK	Thornton's recording of "hound dog" is listed as one of the rock and roll hall of fame's
		KS	Regarded as one of the most significant cultural icons of the 20th century, he is often referred to as the "king of rock and ..."
		DG	Yes, I love him! He is regarded as the king of rock and roll.
Case 2	Context		I love to go bowling with my family, but i'm a horrible bowler. Do you like it?
	Human	KS	Bowling is a sport or leisure activity in which a player rolls or throws a bowling ball towards a target.
		DG	Oh, yes, I love bowling . Rolling that ball down the lane and knocking down the pins gives me a charge.
	SKT	KS	It is one of the major forms of throwing sports.
		RG	I do! I love bowling! I like to play, but i'm not a fan of bowling.
	UKSDG _{PF}	OK	In pin bowling variations, the target is usually to knock over pins at the end of a lane .
		KS	Bowling is a sport or leisure activity in which a player rolls or throws a bowling ball towards a target.
		DG	I love bowling! It is a sport or leisure activity in which a player rolls a bowling ball towards a target.
Case 3	Context		Have you ever seen the Walking Dead?
	Human	KS	The Walking Dead is an American post-apocalyptic horror television series developed by Frank Darabont for AMC that ...
		DG	The American post-apocalyptic series? The one based on the comic book series by Robert Kirkman ? Yes I love it! How about you?
	SKT	KS	The Walking Dead is an American post-apocalyptic horror television series developed by Frank Darabont for AMC that ...
		RG	Yes, I have seen the Walking Dead.
	UKSDG _{PF}	OK	It is based on the comic book series "The Walking Dead" by Robert Kirkman , ...
		KS	The Walking Dead is an American post-apocalyptic horror television series developed by Frank Darabont for AMC that ...
		DG	Yes, I have seen the Walking Dead. It was developed by Frank Darabont for AMC.

Table 5: Examples of generated responses on WoW. KS, DG and OK denote Knowledge Selection, Dialogue Generation and Oracle Knowledge, respectively.

et al., 2019; Kim et al., 2020a; Meng et al., 2020; Chen et al., 2020b; Zheng et al., 2020; Huang et al., 2020). In this paper, we focus on knowledge selection in the unsupervised setting where there is no gold knowledge label. Lian et al. (2019) and Kim et al. (2020a) attempt to deal with this problem via latent models, yet their performance is less than satisfactory. Differently, we design our DDSL to make knowledge selection work well in the unsupervised setting. There is a very recent work (Zhao et al., 2020b) that finetunes GPT-2 (Radford et al., 2019) with the unsupervised pretrained knowledge selection module on unlabeled dialogues. We are different in two aspects: (1) Our DDSL leverages knowledge distillation to alleviate the label noise at the pretraining stage; (2) We adopt the sample weighting idea at the finetuning stage. And we will leverage GPT-2 for future study.

Our work is inspired by Distant Supervision (DS), an effective method to generate labeled data with external knowledge base (KB) for information extraction (Mintz et al., 2009; Min et al., 2013; Zeng et al., 2015; Wang et al., 2018). Following this idea, Gopalakrishnan et al. (2019) use the oracle knowledge from DS to construct the Topical-Chat dataset. Similarly, Qin et al. (2019b) obtain the weakly labeled data to train a KB retriever in

the task-oriented dialogue system. Ren et al. (2019) propose a distantly supervised learning schema at segment level to effectively learn the topic transition vector. Although inspired by the similar idea, we are devoted to knowledge selection in the unsupervised setting, which is a different application of DS. Moreover, rather than just using distant supervision, we design our DDSL with label weighting and knowledge distillation to deal with the noisy labeling problem from DS.

6 Conclusion

We study unsupervised knowledge selection for dialogue generation where the gold knowledge label is not available. Actually, we design the Distilled Distant Supervision Loss, a novel and effective solution to train knowledge selection well in the unsupervised setting. Furthermore, we propose the pretraining-finetuning strategy to deal with the mismatch knowledge selection problem that models tend to select the mismatched knowledge for dialogue generation in the unsupervised setting and will cause the degeneration of knowledge-aware decoder. Experiments show that our approach manages to select knowledge more accurately in the unsupervised setting and generates more informative responses than many strong supervised baselines.

References

- Hengyi Cai, Hongshen Chen, Yonghao Song, Cheng Zhang, Xiaofang Zhao, and Dawei Yin. 2020. [Data Manipulation: Towards Effective Instance Learning for Neural Dialogue Generation via Learning to Augment and Reweight](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6334–6343, Online. Association for Computational Linguistics.
- Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, et al. 2018. Universal Sentence Encoder. *arXiv preprint arXiv:1803.11175*.
- Feilong Chen, Fandong Meng, Jiaming Xu, Peng Li, Bo Xu, and Jie Zhou. 2020a. [DMRM: A dual-channel multi-hop reasoning model for visual dialog](#). In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 7504–7511. AAAI Press.
- Xiuyi Chen, Fandong Meng, Peng Li, Feilong Chen, Bo Xu, Shuang Xu, and Jie Zhou. 2020b. Bridging the Gap between Prior and Posterior Knowledge Selection for Knowledge-Grounded Dialogue Generation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*.
- Xiuyi Chen, Jiaming Xu, and Bo Xu. 2019. [A Working Memory Model for Task-oriented Dialog Response Generation](#). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2687–2693, Florence, Italy. Association for Computational Linguistics.
- Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186.
- Emily Dinan, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston. 2019. Wizard of Wikipedia: Knowledge-Powered Conversational Agents. In *International Conference on Learning Representations*.
- Jiahua Dong, Yang Cong, Gan Sun, Bineng Zhong, and Xiaowei Xu. 2020. What can be transferred: Unsupervised domain adaptation for endoscopic lesions segmentation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4022–4031.
- Marjan Ghazvininejad, Chris Brockett, Ming-Wei Chang, Bill Dolan, Jianfeng Gao, Wen-tau Yih, and Michel Galley. 2018. A knowledge-grounded neural conversation model. In *Thirty-Second AAAI Conference on Artificial Intelligence*.
- Karthik Gopalakrishnan, Behnam Hedayatnia, Qinalang Chen, Anna Gottardi, Sanjeev Kwatra, Anu Venkatesh, Raefer Gabriel, and Dilek Hakkani-Tür. 2019. Topical-Chat: Towards Knowledge-Grounded Open-Domain Conversations. *Proc. Interspeech 2019*, pages 1891–1895.
- Jianping Gou, Baosheng Yu, Stephen John Maybank, and Dacheng Tao. 2020. [Knowledge Distillation: A Survey](#). *CoRR*, abs/2006.05525.
- Jiatao Gu, Zhengdong Lu, Hang Li, and Victor O.K. Li. 2016. [Incorporating Copying Mechanism in Sequence-to-Sequence Learning](#). In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1631–1640, Berlin, Germany. Association for Computational Linguistics.
- Minlie Huang, Xiaoyan Zhu, and Jianfeng Gao. 2020. [Challenges in Building Intelligent Open-Domain Dialog Systems](#). *ACM Trans. Inf. Syst.*, 38(3).
- Byeongchang Kim, Jaewoo Ahn, and Gunhee Kim. 2020a. Sequential Latent Knowledge Selection for Knowledge-Grounded Dialogue. *arXiv: Computation and Language*.
- Seokhwan Kim, Mihail Eric, Karthik Gopalakrishnan, Behnam Hedayatnia, Yang Liu, and Dilek Hakkani-Tür. 2020b. Beyond Domain APIs: Task-oriented Conversational Modeling with Unstructured Knowledge Access. *arXiv preprint arXiv:2006.03533*.
- Diederik P Kingma and Jimmy Ba. 2014. Adam: A Method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980*.
- Philipp Koehn. 2004. Statistical significance tests for machine translation evaluation. In *Proceedings of the 2004 conference on empirical methods in natural language processing*, pages 388–395.
- Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. 2016. [A Diversity-Promoting Objective Function for Neural Conversation Models](#). In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 110–119, San Diego, California. Association for Computational Linguistics.

- Linxiao Li, Can Xu, Wei Wu, Yufan Zhao, Xueliang Zhao, and Chongyang Tao. 2020. [Zero-Resource Knowledge-Grounded Dialogue Generation](#). *CoRR*, abs/2008.12918.
- Margaret Li, Jason Weston, and Stephen Roller. 2019a. ACUTE-EVAL: Improved Dialogue Evaluation with Optimized Questions and Multi-turn Comparisons. *arXiv preprint arXiv:1909.03087*.
- Yuncheng Li, Jianchao Yang, Yale Song, Liangliang Cao, Jiebo Luo, and Li-Jia Li. 2017. Learning from Noisy Labels with Distillation. *2017 IEEE International Conference on Computer Vision (ICCV)*.
- Zekang Li, Cheng Niu, Fandong Meng, Yang Feng, Qian Li, and Jie Zhou. 2019b. Incremental Transformer with Deliberation Decoder for Document Grounded Conversations. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 12–21.
- Rongzhong Lian, Min Xie, Fan Wang, Jinhua Peng, and Hua Wu. 2019. Learning to Select Knowledge for Response Generation in Dialog Systems. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 5081–5087. AAAI Press.
- Xiexiong Lin, Weiyu Jian, Jianshan He, Taifeng Wang, and Wei Chu. 2020. [Generating Informative Conversational Response using Recurrent Knowledge-Interaction and Knowledge-Copy](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 41–52, Online. Association for Computational Linguistics.
- Shuman Liu, Hongshen Chen, Zhaochun Ren, Yang Feng, Qun Liu, and Dawei Yin. 2018. [Knowledge Diffusion for Neural Dialogue Generation](#). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1489–1498, Melbourne, Australia. Association for Computational Linguistics.
- Chuan Meng, Pengjie Ren, Zhumin Chen, Christof Monz, Jun Ma, and Maarten de Rijke. 2019. RefNet: A Reference-aware Network for Background Based Conversation. *arXiv preprint arXiv:1908.06449*.
- Chuan Meng, Pengjie Ren, Zhumin Chen, Zhaochun Ren, Tengxiao Xi, and Maarten de Rijke. 2021. Initiative-Aware Self-Supervised learning for Knowledge-Grounded Conversations. In *Proceedings of the 44rd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '21*. Association for Computing Machinery.
- Chuan Meng, Pengjie Ren, Zhumin Chen, Weiwei Sun, Zhaochun Ren, Zhaopeng Tu, and Maarten de Rijke. 2020. [DukeNet: A Dual Knowledge Interaction Network for Knowledge-Grounded Conversation](#). In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '20*, pages 1151–1160, New York, NY, USA. Association for Computing Machinery.
- Xiang Gao Bill Dolan Jianfeng Gao Michel Galley, Chris Brockett. 2018. Grounded Response Generation Task at DSTC7. *Proceedings of the AAAI-19 Workshop on Dialog System Technology Challenges*.
- Bonan Min, Ralph Grishman, Li Wan, Chang Wang, and David Gondek. 2013. [Distant Supervision for Relation Extraction with an Incomplete Knowledge Base](#). In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 777–782, Atlanta, Georgia. Association for Computational Linguistics.
- Mike Mintz, Steven Bills, Rion Snow, and Daniel Jurafsky. 2009. [Distant supervision for relation extraction without labeled data](#). In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 1003–1011, Suntec, Singapore. Association for Computational Linguistics.
- Nikita Moghe, Siddhartha Arora, Suman Banerjee, and Mitesh M Khapra. 2018. Towards Exploiting Background Knowledge for Building Conversation Systems. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2322–2332.
- Jinjie Ni, Tom Young, Vlad Pandealea, Fuzhao Xue, Vinay Adiga, and Erik Cambria. 2021. [Recent advances in deep learning-based dialogue systems](#). *CoRR*, abs/2105.04387.
- Zheng-Yu Niu, Hua Wu, Haifeng Wang, et al. 2019. Knowledge Aware Conversation Generation with Explainable Reasoning over Augmented Graphs. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 1782–1792.
- Gabriel Pereyra, George Tucker, Jan Chorowski, Łukasz Kaiser, and Geoffrey Hinton. 2017. Regularizing Neural Networks by Penalizing Confident Output Distributions. *arXiv preprint arXiv:1701.06548*.
- Lianhui Qin, Michel Galley, Chris Brockett, Xiaodong Liu, Xiang Gao, Bill Dolan, Yejin Choi, and Jianfeng Gao. 2019a. Conversing by Reading: Contentful Neural Conversation with On-demand Machine Reading. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5427–5436.
- Libo Qin, Yijia Liu, Wanxiang Che, Haoyang Wen, Yangming Li, and Ting Liu. 2019b. [Entity-Consistent End-to-end Task-Oriented Dialogue System with KB Retriever](#). In *Proceedings of the*

- 2019 *Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 133–142, Hong Kong, China. Association for Computational Linguistics.
- Alec Radford, Jeff Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. Language Models are Unsupervised Multitask Learners.
- Pengjie Ren, Zhumin Chen, Christof Monz, Jun Ma, and Maarten de Rijke. 2019. Thinking Globally, Acting Locally: Distantly Supervised Global-to-Local Knowledge Selection for Background Based Conversation. *arXiv preprint arXiv:1908.09528*.
- Bharat Bhushan Sau and Vineeth N Balasubramanian. 2016. Deep Model Compression: Distilling Knowledge from Noisy Teachers. *arXiv preprint arXiv:1610.09650*.
- C. Sellitz, L.S. Wrightsman, S.W. Cook, and Society for the Psychological Study of Social Issues. 1976. *Research Methods in Social Relations*. Holt, Rinehart and Winston.
- Yajing Sun, Yue Hu, Luxi Xing, Jing Yu, and Yuqiang Xie. 2020. History-adaption Knowledge Incorporation Mechanism for Multi-turn Dialogue System. In *AAAI*.
- Xiangru Tang and Po Hu. 2019. Knowledge-Aware Self-Attention Networks for Document Grounded Dialogue Generation. In *International Conference on Knowledge Science, Engineering and Management*, pages 400–411. Springer.
- Zhiliang Tian, Wei Bi, Dongkyu Lee, Lanqing Xue, Yiping Song, Xiaojiang Liu, and Nevin L. Zhang. 2020. [Response-Anticipated Memory for On-Demand Knowledge Integration in Response Generation](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 650–659, Online. Association for Computational Linguistics.
- Yi-Lin Tuan, Wei Wei, and William Yang Wang. 2020. Unsupervised Injection of Knowledge into Dialogue Generation via Language Models. *arXiv preprint arXiv:2004.14614*.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008.
- Guanying Wang, Wen Zhang, Ruoxu Wang, Yalin Zhou, Xi Chen, Wei Zhang, Hai Zhu, and Huajun Chen. 2018. [Label-Free Distant Supervision for Relation Extraction via Knowledge Graph Embedding](#). In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2246–2255, Brussels, Belgium. Association for Computational Linguistics.
- Wenquan Wu, Zhen Guo, Xiangyang Zhou, Hua Wu, Xiyuan Zhang, Rongzhong Lian, and Haifeng Wang. 2019. [Proactive Human-Machine Conversation with Explicit Conversation Goal](#). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3794–3804, Florence, Italy. Association for Computational Linguistics.
- Yingce Xia, Fei Tian, Lijun Wu, Jianxin Lin, Tao Qin, Nenghai Yu, and Tie-Yan Liu. 2017. Deliberation Networks: Sequence Generation Beyond One-Pass Decoding. In *Advances in Neural Information Processing Systems*, pages 1784–1794.
- Qizhe Xie, Minh-Thang Luong, Eduard Hovy, and Quoc V Le. 2020. Self-training with Noisy Student improves ImageNet classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10687–10698.
- Yan Xu, Etsuko Ishii, Zihan Liu, Genta Indra Winata, Dan Su, Andrea Madotto, and Pascale Fung. 2021. [Retrieval-free knowledge-grounded dialogue response generation with adapters](#). *CoRR*, abs/2105.06232.
- Kaijia Yang, Liang He, Xin-yu Dai, Shujian Huang, and Jiajun Chen. 2019a. [Exploiting Noisy Data in Distant Supervision Relation Classification](#). In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 3216–3225, Minneapolis, Minnesota. Association for Computational Linguistics.
- Liu Yang, Junjie Hu, Minghui Qiu, Chen Qu, Jianfeng Gao, W. Bruce Croft, Xiaodong Liu, Yelong Shen, and Jingjing Liu. 2019b. [A Hybrid Retrieval-Generation Neural Conversation Model](#). In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management, CIKM 2019, Beijing, China, November 3-7, 2019*, pages 1341–1350. ACM.
- Semih Yavuz, Abhinav Rastogi, Guan-Lin Chao, Dilek Hakkani-Tür, and Amazon Alexa AI. 2019. DEEP-COPY: Grounded Response Generation with Hierarchical Pointer Networks. In *20th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, page 122.
- Hao-Tong Ye, Kai-Lin Lo, Shang-Yu Su, and Yun-Nung Chen. 2020. Knowledge-grounded response generation with deep attentional latent-variable model. *Computer Speech & Language*, page 101069.
- Wenhao Yu, Chenguang Zhu, Zaitang Li, Zhiting Hu, Qingyun Wang, Heng Ji, and Meng Jiang. 2020. [A survey of knowledge-enhanced text generation](#). *CoRR*, abs/2010.04389.
- Li Yuan, Francis EH Tay, Guilin Li, Tao Wang, and Jiashi Feng. 2020. Revisiting Knowledge Distilla-

- tion via Label Smoothing Regularization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3903–3911.
- Daojian Zeng, Kang Liu, Yubo Chen, and Jun Zhao. 2015. [Distant Supervision for Relation Extraction via Piecewise Convolutional Neural Networks](#). In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1753–1762, Lisbon, Portugal. Association for Computational Linguistics.
- Duzhen Zhang, Xiuyi Chen, Shuang Xu, and Bo Xu. 2020. Knowledge aware emotion recognition in textual conversations via multi-task incremental transformer. In *COLING*.
- Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. 2018. Personalizing Dialogue Agents: I have a dog, do you have pets too? In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2204–2213.
- Yangjun Zhang, Pengjie Ren, and Maarten de Rijke. 2019. Improving background based conversation with context-aware knowledge pre-selection. *arXiv preprint arXiv:1906.06685*.
- Xueliang Zhao, Wei Wu, Chongyang Tao, Can Xu, Dongyan Zhao, and Rui Yan. 2020a. [Low-Resource Knowledge-Grounded Dialogue Generation](#). In *International Conference on Learning Representations*.
- Xueliang Zhao, Wei Wu, Can Xu, Chongyang Tao, Dongyan Zhao, and Rui Yan. 2020b. Knowledge-Grounded Dialogue Generation with Pre-trained Language Models. *arXiv preprint arXiv:2010.08824*.
- Chujie Zheng, Yunbo Cao, Daxin Jiang, and Minlie Huang. 2020. [Difference-aware Knowledge Selection for Knowledge-grounded Conversation Generation](#). *CoRR*, abs/2009.09378.
- Wen Zheng and Ke Zhou. 2019. Enhancing Conversational Dialogue Models with Grounded Knowledge. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pages 709–718.
- Hao Zhou, Chujie Zheng, Kaili Huang, Minlie Huang, and Xiaoyan Zhu. 2020. [KdConv: A Chinese Multi-domain Dialogue Dataset Towards Multi-turn Knowledge-driven Conversation](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7098–7108, Online. Association for Computational Linguistics.
- Kangyan Zhou, Shrimai Prabhumoye, and Alan W Black. 2018. A Dataset for Document Grounded Conversations. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 708–713.
- Wenya Zhu, Kaixiang Mo, Yu Zhang, Zhangbin Zhu, Xuezheng Peng, and Qiang Yang. 2017. Flexible End-to-End Dialogue System for Knowledge Grounded Conversation. *arXiv preprint arXiv:1709.04264*.

A Additional Experiments

A.1 Models for Comparison

A.1.1 Our Unsupervised Methods

We implement our model in the unsupervised setting, namely Unsupervised Knowledge Selection for Dialogue Generation (**UKSDG**), which is optimized with our DDSL in Equation 11 for unsupervised knowledge selection and generation loss in Equation 7. **UKSDG_{PF}** indicates that we adopt our Pretraining-Finetuning strategy to alleviate the mismatch knowledge selection problem in Section 2.7. Furthermore, we remove several components for ablation study:

- (1) **UKSDG w/o DDSL** does not use our DDSL.
- (2) **UKSDG_{vec} w/o DDSL** further replaces the decoder input \mathbf{H}_{rc} (in Section 2.5) with the averaged knowledge vector enhanced context $\mathbf{H}_{vec} = \mathbf{H}_{x_t} + \sum_{k_t^i \in \mathcal{K}_t} \mathcal{S}_{k_t^i} \cdot \mathbf{h}_{k_t^i}$.
- (3) **UKSDG_{PF} w/o SW** does not use the Sample Weighting in Equation 13.

A.1.2 Our Supervised Methods

Here, we also report the Knowledge Selection for Dialogue Generation (**KSDG**) in the supervised setting as described in Section 2.3, 2.4 and 2.5. And **KSDG_{PF}** alleviates the mismatch problem using the pretrain-finetuning strategy in Section 2.7. We use the gold knowledge loss in Equation 5 to train KSDG and KSDG_{PF} in the supervised setting while we train UKSDG and UKSDG_{PF} with our DDSL when the gold knowledge is not available.

We compare our method with a set of baselines:

A.1.3 No Knowledge

S2S_{Transformer} is a Seq2Seq model based on Transformer (Vaswani et al., 2017) that does not leverage the external knowledge.

S2S_{BERT} replaces the Transformer Encoder with a pre-trained BERT (Devlin et al., 2019).

KnowExpert (Xu et al., 2021) avoids the knowledge retrieval process and attempts to inject prior knowledge into the pre-trained language models for knowledge-grounded dialogue generation task. Essentially, KnowExpert stores knowledge in its parameters with lightweight adapters. Therefore, KnowExpert does not use knowledge explicitly at inference.

A.1.4 Supervised Knowledge Selection

TMN is short for End-to-End Transformer Mem-Net (Dinan et al., 2019), which selects knowledge based on the Transformer memory network and generate responses via the Transformer decoder.

TMN_{BERT+PostKS+CP}, implemented by (Kim et al., 2020a), enhances the encoder with the pretrained BERT, knowledge selection module with PostKS (Lian et al., 2019) and decoder with the copy mechanism (CP).

SKT (Kim et al., 2020a), short for Sequential Knowledge Transformer, uses the posterior distribution by sequential latent modeling and achieves promising results in the supervised setting.

SKT+PIPM+KDBTS (Chen et al., 2020b) equips SKT with Posterior Information Prediction Module (PIPM) and Knowledge Distillation Based Training Strategy (KDBTS) to bridge the gap between prior and posterior knowledge selection for knowledge-grounded dialogue generation.

A.1.5 Unsupervised Knowledge Selection

TMN₀ is **TMN**, trained only via generation loss in Equation 7 without knowledge loss in Equation 5.

SKT₀ is **SKT** optimized without knowledge loss.

PostKS (Lian et al., 2019) takes the benefits of latent variable models and leverages the posterior knowledge distribution as a pseudo label for knowledge selection without knowledge loss. Here, we use the results provided by (Kim et al., 2020a).

A.1.6 Knowledge-Aware Generation

Different from ours on the knowledge selection setting, there are much work, named knowledge-aware generation (Huang et al., 2020), which regard the retrieved knowledge pool as the pseudo pre-identified document for dialogue generation with complex knowledge injection mechanism. Nevertheless, we provide the results to get a comprehensive understanding where this field is going. **MTASK-RF** (Ghazvininejad et al., 2018) is an early model that realizes knowledge-grounded conversation without crowd-sourced knowledge-grounded dialogues. Here we use the results provided by (Li et al., 2020).

ITDD is short for Incremental Transformer with Deliberation Decoder (Li et al., 2019b) where the encoder incrementally represents multi-turn dialogues and knowledge, and the decoder conducts response decoding in two passes.

DRD is short for Disentangle Response Decoder (Zhao et al., 2020a), a model that exploits pre-

training techniques to tackle the low-resource challenge in knowledge-grounded dialogue generation. We choose the one whose parameters are fine-tuned on the full training data.

KIC (Lin et al., 2020) integrates recurrent Knowledge-Interaction and knowledge Copy (KIC) to generate informative responses.

ZRKGK (Li et al., 2020) is a very recent and unpublished work, which learns their model under the zero-resource setting, where the dialogue corpus and the knowledge corpus that are independent with each other.

A.2 Quantitative Results

Table 6 shows the automatic results in various settings on the Wizard of Wikipedia dataset, from which we have the following observations⁹: (1) The models with retrieved document generally achieves much lower PPL, still the R-1 is worse than our UKSDG_{PF}. For the worse R-1 score, we think the retrieved document contains the redundant and irrelevant information which are harmful for dialogue generation (Zhao et al., 2020b). Meanwhile, we attribute the better PPL score to the complex knowledge injection mechanism. For example, ITDD leverages the deliberation decoder (Xia et al., 2017) to improve context coherence and knowledge correctness and DRD devises a disentangled response decoder with pretrained language model, context processor and knowledge processor. (2) Comparing KSDG with KSDG_{PF}, again we see the generation quality improvement via the pretrain-finetuning strategy. Hence, our method is general in both supervised and unsupervised setting. (3) Comparing UKSDG_{vec} w/o DDSL with UKSDG w/o DDSL, we see that although selecting the knowledge vector softly allows the gradient from dialogue generation to update knowledge selection directly, compressing the knowledge into a vector will loss much information for dialogue generation. This observation, combined with observations (1) and (2), together indicates the importance of leveraging the selected knowledge well for future study.

⁹The observations in the main paper will not be discussed again here.

Setting	Method	Test Seen				Test Unseen			
		Acc	PPL	R-1	R-2	Acc	PPL	R-1	R-2
No Knowledge	S2S _{Transformer} [†] (Vaswani et al., 2017)	N/A	41.8	17.8	N/A	N/A	87.0	14.0	N/A
	S2S _{BERT} [†]	N/A	55.2	17.0	4.7	N/A	93.3	15.0	3.1
	KnowExpert (Xu et al., 2021)	N/A	15.3	18.8	N/A	N/A	21.2	16.6	N/A
Retrieved Document	MTASK-RF (Ghazvininejad et al., 2018)	N/A	65.4	13.1	N/A	N/A	67.7	12.3	N/A
	ITDD (Li et al., 2019b)	N/A	17.8	16.2	N/A	N/A	44.8	11.4	N/A
	DRD (Zhao et al., 2020a)	N/A	23.0	18.0	N/A	N/A	25.6	16.5	N/A
	KIC (Lin et al., 2020)	N/A	51.9	18.4	N/A	N/A	65.8	17.3	N/A
	ZRKGK (Li et al., 2020)	N/A	41.1	18.9	N/A	N/A	42.7	18.8	N/A
Unsupervised KS	TMN ₀ (Dinan et al., 2019)	13.4	66.5	15.9	N/A	11.8	103.6	14.3	N/A
	PostKS (Lian et al., 2019)	4.8	79.1	13.0	1.0	4.2	193.8	13.1	1.0
	SKT ₀ (Kim et al., 2020a)	0.3	54.7	17.1	4.6	0.1	88.2	15.5	3.4
	UKSDG w/o DDSL	5.2	49.3	17.4	5.1	5.1	82.8	15.1	3.3
	UKSDG _{vec} w/o DDSL	13.5	46.1	17.8	5.2	13.1	82.3	15.5	3.3
	UKSDG (ours)	23.8	51.8	19.5	6.8	16.2	76.3	16.3	4.4
	UKSDG _{PF} w/o SW (ours)	26.4	44.1	20.3	7.4	20.8	64.5	17.8	5.6
	UKSDG _{PF} (ours)	26.4	45.0	20.6	7.7	20.8	65.5	18.2	5.9
	SKT _{1/2} (Kim et al., 2020a)	25.1	49.0	19.2	6.6	16.7	77.8	16.1	4.1
	SKT _{1/4} (Kim et al., 2020a)	22.4	45.7	18.7	6.1	13.8	78.0	15.8	3.6
Semi-supervised KS	SKT _{1/8} (Kim et al., 2020a)	21.0	45.3	18.6	6.0	12.3	79.9	15.7	3.6
	TMN (Dinan et al., 2019)	21.1	63.5	16.9	NA	14.3	97.3	14.4	NA
	TMN _{E2E} +BERT+PostKS+CP	25.5	52.2	19.0	6.5	14.4	83.4	15.6	3.9
Supervised KS	SKT (Kim et al., 2020a)	26.8	52.0	19.3	6.8	18.3	81.4	16.1	4.2
	SKT+PIPM+KDBTS (Chen et al., 2020b)	27.7	42.7	19.9	7.3	19.4	65.7	17.6	5.4
	KSDG (our)	25.1	53.0	18.9	6.5	17.3	80.8	16.3	4.5
	KSDG _{PF} (our)	28.4	45.1	20.4	7.9	21.2	71.4	17.9	5.6

Table 6: Quantitative results on the Wizard of Wikipedia dataset. Note that models with “†” are implemented by ourselves and other models with citation are from the original paper. KS denotes knowledge selection.