Fine-grained Post-training for Improving Retrieval-based Dialogue Systems

Janghoon Han[1,3], **Taesuk Hong**[1], **Byoungjae Kim**[1], **Youngjoong Ko**[2], **Jungyun Seo**[1]

[1]Department of Computer Science and Engineering, Sogang University
[2]Department of Computer Science and Engineering, Sungkyunkwan University
[3]LG AI Research

`janghoon.han@lgresearch.ai`
`{hongtaesuk,wiz3021,seojy}@sogang.ac.kr`
`yjko@skku.edu`

Abstract

Retrieval-based dialogue systems display an outstanding performance when pre-trained language models are used, which includes bidirectional encoder representations from transformers (BERT). During the multi-turn response selection, BERT focuses on training the relationship between the context with multiple utterances and the response. However, this method of training is insufficient when considering the relations between each utterance in the context. This leads to a problem of not completely understanding the context flow that is required to select a response. To address this issue, we propose a new fine-grained post-training method that reflects the characteristics of the multi-turn dialogue. Specifically, the model learns the utterance level interactions by training every short context-response pair in a dialogue session. Furthermore, by using a new training objective, the utterance relevance classification, the model understands the semantic relevance and coherence between the dialogue utterances. Experimental results show that our model achieves new state-of-the-art with significant margins on three benchmark datasets. This suggests that the fine-grained post-training method is highly effective for the response selection task.[1]

1   Introduction

Constructing a dialogue system that can naturally and consistently interact with humans is currently a popular research topic. There are two approaches for the implementation of a dialogue system: generation-based and retrieval-based methods. The latter approach aims to select the correct response among the response candidates. In the initial multi-turn response selection, Lowe et al. (2015) proposed leveraging RNN to match the dialogue context with a response. Later, with the advent of the attention mechanism (Bahdanau et al., 2015; Luong et al., 2015; Vaswani et al., 2017), multi-turn response selection models that use the attention mechanism have been proposed (Zhou et al., 2018). Recently, pre-trained language models, such as bidirectional encoder representations from transformers, BERT, have been applied to a variety of response selection models (Vig and Ramea, 2019; Lu et al., 2020; Gu et al., 2020), and they have shown excellent performance.

Recently, pre-trained language models have been widely used in several natural language processing areas, such as question answering and dialogue systems. One of the best pre-trained language models, BERT (Devlin et al., 2019), is initially pre-trained on a large and general domain corpus, and then it is fine-tuned to adapt to specific tasks. Since BERT is pre-trained with general data, its performance can be improved by post-training to adapt to domain-specific data. Some previous studies (Xu et al., 2019; Whang et al., 2020) proposed a post-training method that can learn domain data before fine-tuning for a task. In the previous studies, the models were post-trained using domain-specific task data with the same pre-training objectives as BERT, masked language model (MLM) and next sentence prediction (NSP).

To develop a new post-training method that is suitable for dialogue, we propose a simple but powerful fine-grained post-training method. The new post-training method has two learning strategies. The first is to train the model by dividing the entire dialogue into multiple short context–response pairs. The second is to train the model with a new objective called utterance relevance classification, which classifies the relation between given utterances and the target utterance into more fine-grained labels.

The dialogue consists of a context that includes multiple utterances and a response with one utterance. There are two advantages to learning the dialogue by dividing it into multiple new short context-

response pairs, rather than learning with the entire context–response pair during post-training. First, the model can learn the interaction between internal utterances, which is overlooked in the previous training methods. The previous multi-turn response selection models focus on identifying the associated information between a context with multiple utterances and the response. To understand the associated information, BERT takes the whole context as input to represent the relationship between the context and the response, instead of gradually expanding and learning the relationship between the utterances inside the context. The relationship between the entire context and response can be learned through self-attention. However, the relationship between the utterances in the dialogue is easily overlooked. To address this issue, we divide the entire dialogue into multiple short context–response pairs. Since each pair consists of internal utterances, the model can learn the utterance-level interactions. The second advantage is that the model can capture the relationship between the utterances more accurately. In general, the utterances that are related to the response are located close to the response. As short context-response pairs consist only of utterances that are close to the response, more fine-grained training is possible.

Another strategy of fine-grained post-training is that it involves using a new training objective that is called the utterance relevance classification (URC). In the case of the NSP used in BERT, the model distinguishes whether the target utterance is random or the next. As mentioned by Lan et al. (2020), the model trained with the NSP can easily learn the topic prediction that distinguishes the semantic meaning of the utterances. However, it lacks coherence prediction that distinguishes whether the selected utterance is consecutive. In the case of sentence ordering prediction (SOP) used in Lan et al. (2020), the coherence between the utterances is well learned because the order of the two sequences is trained. However, topic prediction is relatively insufficient because the two sequences are semantically similar. As it is important to distinguish between semantically similar utterances in the multi-turn dialogue and determine whether the selected utterances are consecutive, we propose URC, which classifies the target utterance into three categories (random, semantically similar, next) to learn the topics and coherence.

The contributions of our study are summarized as follows:

1. Through short context-response pair training during fine-grained post-training, the model effectively learns the interactions between internal utterances, which can be easily overlooked in the existing methods. This significantly improves the performance of response selection.

2. By devising the new training objective, URC, we enhance the model's capability to measure both the semantic relevance and coherence between utterances, improving the model to select the appropriate response.

We achieved state-of-the-art performance with a significant improvement for three benchmarks (Ubuntu, Douban, E-commerce). Specifically, our model achieved an absolute improvement in $R_{10}@1$ by 2.7%p, 0.6%p, and 9.4%p on Ubuntu Corpus V1, Douban Corpus, and E-commerce Corpus, respectively, in comparison to previous state-of-the-art methods. The results indicate the effectiveness and generality of the proposed method.

2 Related Work

The existing methods for building dialogue systems can be categorized into two groups: those with a retrieval-based approach (Chaudhuri et al., 2018; Tao et al., 2019; Yuan et al., 2019) and those with a generation-based approach (Wu et al., 2018; Zhou et al., 2018; Hosseini-Asl et al., 2020; Ham et al., 2020). Recent studies have focused on the multi-turn retrieval dialogue system where the system selects the most appropriate response when a multi-turn dialogue context is provided. Lowe et al. (2015) proposed a new benchmark dataset called the Ubuntu internet relay chat (IRC) Corpus V1 and a RNN-based baseline model. Kadlec et al. (2015) suggested a dual encoder-based model that attempts to effectively encode the context and response by using LSTM and CNN as encoder. With the advent of the attention mechanism (Bahdanau et al., 2015; Luong et al., 2015; Vaswani et al., 2017), models such as the deep attention matching network (Zhou et al., 2018), which applied the attention mechanism to the response selection dialogue system, have been proposed. Chen and Wang (2019) adapted the natural language inference model to the response selection task. Tao et al. (2019) performed a deep interaction between

the context and the response through multiple interaction blocks. Yuan et al. (2019) improved the performance by controlling the dialogue context information with a multi-hop selector.

The pre-trained language models have shown an impressive performance in the response selection (Lu et al., 2020; Gu et al., 2020; Whang et al., 2021; Xu et al., 2021). One of those, BERT, is a bidirectional transformer-based encoder that has multiple layers. We use the publicly opened BERT$_{base}$ model in which the number of layers, attention head, and size of the hidden state are 12, 12, and 768, respectively.

There are a variety of training objectives for the pre-trained language models. BERT uses two training objectives: MLM and NSP. The former randomly masks 15% of the tokens that are predicted by the model. This method of training aims for the model to learn the overall contextual representation of a given text. In the latter method, the model is given two sequences of text: A and B. The model is trained to determine if sequence B is the next sequence after sequence A. The model takes the input, sequences A and B, separated by the special token SEP. The model uses the segment embeddings of 0 for sequence A and 1 for sequence B. Then, by using the CLS token, the model predicts the relationship between sequences A and B. AL-BERT (Lan et al., 2020) uses sentence ordering prediction (SOP) instead of NSP as the training objectives. The SOP distinguishes whether the order of sequences A and B is correct or if they have been swapped.

The post-training method, which helps the model understand a certain domain, was introduced in the response selection task (Whang et al., 2020; Gu et al., 2020; Humeau et al., 2020; Whang et al., 2021; Xu et al., 2021). In addition to domain adaptation, the post-training method has the advantage of data augmentation because it learns the relationship between the two sequences in the dialogue session with the NSP. However, the method does not reflect the conversational characteristics because it merely follows BERT's pre-training method. To address this issue, we propose a novel post-training method that is suitable for a multi-turn dialogue. The proposed method achieved better performance in comparison to the previous post-training.

3 Model

3.1 Problem Formalization

Suppose that the dataset $D = \{(c_i, r_i, y_i)\}_{i=1}^N$ is a set of N triples that consist of the context c_i, response r_i, and ground truth label y_i. The context is a sequence of utterances, which is $c_i = \{u_1, u_2, ..., u_M\}$, where M is the maximum context length. The j_{th} utterance $u_j = \{w_{j,1}, w_{j,2}, ..., w_{j,L}\}$ contains L tokens, where L is the maximum sequence length. Each response, r_i, is a single utterance. $y_i \in \{0, 1\}$ denotes the truth label of a given triple where $y_i = 1$. This indicates that r_i is the correct response for the context c_i; otherwise, $y_i = 0$. The task is to find the matching model, $g(\cdot, \cdot)$, for the D. The matching degree of c_i and r_i is obtained through $g(c_i, r_i)$ for a given context–response pair (c_i, r_i).

3.2 Fine-tuning BERT for Response Selection

This study is based on the binary classification to fine-tune BERT for the response selection task that analyzes the relationship between the context and response. The input format (x) of the existing BERT model is ($[CLS]$, $sequenceA$, $[SEP]$, $sequenceB$, $[SEP]$), where $[CLS]$ and $[SEP]$ are CLS and SEP tokens, respectively.

To measure the matching degree of a context-response pair, we construct the input by using sequence A as a context and sequence B as a response. In addition, the end of the utterance token (EOU) is placed at the end of each utterance to distinguish them in the context. The input format of BERT for the response selection is as follows:

$$x = [CLS]\ u_1\ [EOU]\ ...\ u_M\ [EOU]\ [SEP] \\ r_i\ [SEP] \quad (1)$$

x subsequently becomes input representation vectors through the sum of the position, segment, and token embedding. The transformer block in BERT calculates the cross attention between the input representation of the context and the response through the self-attention mechanism. Then, the final hidden vector of the first input token in BERT, $T_{[CLS]}$, is used as the aggregate representation of the context–response pair. The final score $g(c, r)$, which is the matching degree between the context and the response, is obtained by passing $T_{[CLS]}$ through a single-layer neural network.

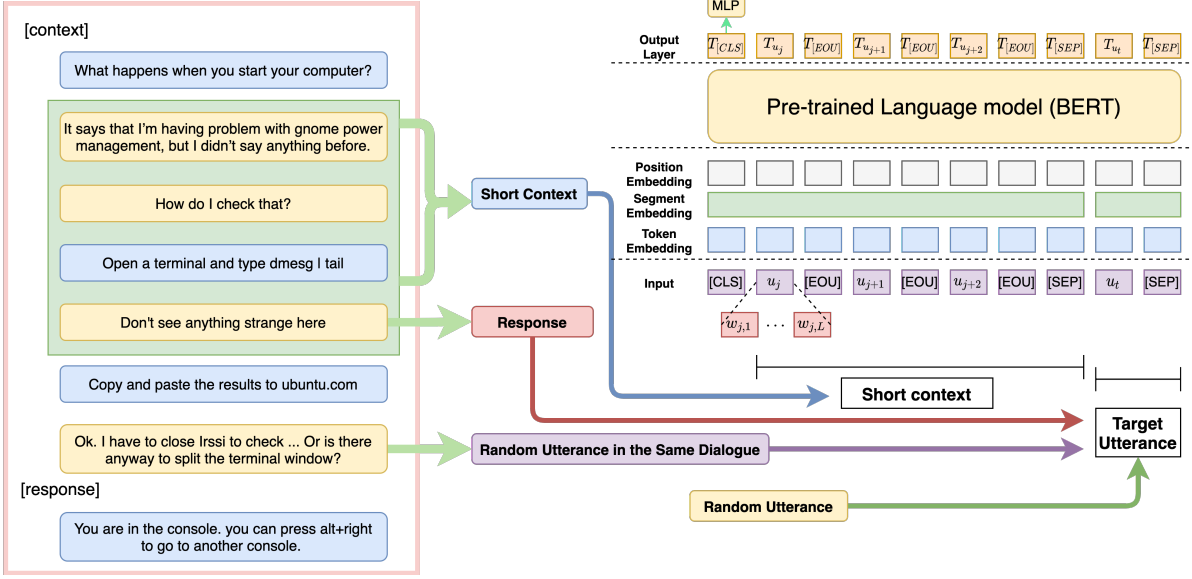$$g(c, r) = \sigma(W_{fine}T_{[CLS]} + b) \quad (2)$$

Figure 1: Architecture of fine-grained post-training. The short context length k is three.

where W_{fine} is a task-specific trainable parameter for fine-tuning. Eventually, the weights of the model are updated by using the cross-entropy loss function.

$$Loss = -\sum_{(c_i,r_i,y_i)\in D} y_i \log(g(c_i, r_i))$$
$$+ (1 - y_i) \log(1 - g(c_i, r_i)) \quad (3)$$

3.3 Fine-grained Post-training

To improve the capability of selecting an appropriate response by effectively grasping multi-turn dialogue information, we propose a simple but powerful fine-grained post-training method in Figure 1. The fine-grained post-training method has two learning strategies. The entire dialogue session is divided into multiple short context-response pairs, and URC is used as one of the training objectives. Through the former strategy, the model learns the interaction of the related internal utterances of the dialogue. Through URC, it learns the semantic relevance and coherence between the utterances.

3.3.1 Short Context-response Pair Training

We post-train the model by constructing multiple short context-response pairs using all utterances of the dialogue session to learn the utterance level interaction. We regard every utterance as a response and its previous k utterances as a short context. The short context contains fewer utterances than the average number of utterances in the dialogue sessions.
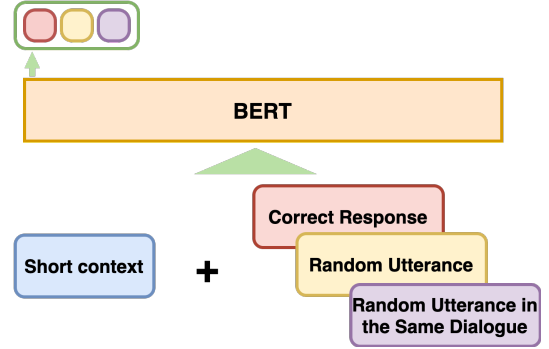

Figure 2: Utterance relevance classification of the fine-grained post-training method

Each short context-response pair is trained to learn the internal utterance interactions, eventually allowing the model to understand the relationship between all the utterances in a dialogue session. It also allows the model to learn the interaction of the utterances closely related to the response because the context is appropriately configured with a short length.

3.3.2 Utterance Relevance Classification

The NSP objective (Devlin et al., 2019) is inadequate for capturing the coherence between the utterances. This is because NSP mainly learns the topic's semantic relevance by classifying between a random and the next utterance. By using the SOP (Lan et al., 2020) as an objective function, the ability to distinguish the semantic relevance decreases

1552

because the model learns the coherence of two utterances with a similar topic. To learn both the semantic relevance and the coherence in a dialogue, we propose a new training objective that is called the utterance relevance classification (URC) in Figure 2. The URC classifies the target utterance for a given short context into one of three labels. The first label is a random utterance. Secondly, an utterance, which is not the response, is randomly sampled in the same dialogue session. Although utterances of the same dialogue session have a similar topic to the correct response, they are inappropriate for the coherence prediction. Finally, the correct response is selected. The model learns the topic prediction by performing a classification between the random utterances and correct responses, and the model makes the coherence predictions by classifying the random utterances and correct responses in the same dialogue sessions. By classifying the relationship between the short context and the target utterance into three cases, the model can learn both the semantic relevance information and the coherence information of the dialogue session.

3.3.3 Training Setup

An overview of the fine-grained post-training (FP) method is shown in Figure 1. First, when given the conversation session $U_i = \{u_1, u_2, ..., u_M, u_{M+1} = r_i\}$, we select the continuous utterances and form a short context–response pair $S_j = \{u_j, u_{j+1}, ..., u_{j+k-1}, u_{j+k}\}$ with a context length of k. The model classifies the relationship between a short context $sc = \{u_j, u_{j+1}, ..., u_{j+k-1}\}$ and the given target utterance u_t. The target utterance can be one of three options: a random utterance u_r, a random utterance for the same dialogue session u_s, or the response u_{j+k}, where $1 \leq s \leq M+1$ and $j + k \neq s$. We denote the input sequence x for the fine-grained post-training as follows:

$$x = [CLS]\ u_j\ [EOU]\ ...\ u_{j+k-1}\ [EOU]$$
$$[SEP]\ u_t\ [SEP] \quad (4)$$

As an aggregate representation, $T_{[CLS]}$ is used. The final score $g_{urc}(sc, u_t)$ is obtained by feeding $T_{[CLS]}$ through a single-layer perceptron, and the degree of relevance between the short context and target utterance is obtained through the score. To calculate the URC loss, we use the cross-entropy loss, which is formulated as follows:

$$L_{URC} = -\sum \sum_i^3 y_i \log(g_{urc}(sc, u_t)_i) \quad (5)$$

To train the proposed model, we use the MLM and URC together. In the case of the MLM, we apply a dynamic masking technique proposed by RoBERTa (Liu et al., 2019), which is unlike BERT. The model can learn more contextual representations because it learns by masking a random token each time instead of learning by masking a predetermined token. To optimize the model, we use the sum of the cross-entropy loss of the MLM and URC, which is formulated as follows:

$$L_{FP} = L_{MLM} + L_{URC} \quad (6)$$

4 Experiments

4.1 Datasets

We tested our model on widely used benchmarks that include Ubuntu Corpus V1, Douban Corpus, and the E-commerce Corpus. The statistics for the three datasets are presented in Table 1.

- **Ubuntu Corpus**
  The Ubuntu IRC Corpus V1 (Lowe et al., 2015) is chatting log conversations, a publicly available domain-specific dialogue dataset. This dialogue data deals with Ubuntu-related topics. In our study, the data proposed by Xu et al. (2017) are used. The data are preprocessed with special placeholders such as numbers, URLs, and system paths.

- **Douban Corpus**
  Douban Corpus (Wu et al., 2017) is a Chinese open-domain dataset from the Douban group, which is a popular social networking service. It consists of dyadic dialogues (i.e., a conversation between two people) that is longer than two turns.

- **E-commerce Corpus**
  The E-commerce Corpus (Zhang et al., 2018) is a Chinese multi-turn dialogue that is collected from Taobao, which is the largest e-commerce platform in China. It contains real-world conversations between customers and customer service staff. The corpus consists of diverse conversations such as consultations and recommendations.

Dataset	Ubuntu			Douban			E-commerce		
	Train	Val	Test	Train	Val	Test	Train	Val	Test
# context-response pairs	1M	500K	500K	1M	50K	50K	1M	10K	10K
# candidate per context	2	10	10	2	2	10	2	2	10
# positive : negative	1:1	1:9	1:9	1:1	1:1	1.2:8.8	1:1	1:1	1:9
# avg turns	10.13	10.11	10.11	6.69	6.75	6.45	5.51	5.48	5.64

Table 1: Data statistics for the Ubuntu, Douban, and E-Commerce Corpus

4.2 Post-training Data

For the fine-grained post-training, we reconstructed the three benchmark datasets. Specifically, out of the one million triples in each benchmark's training set, we used 500K positive triples as dialogue sessions. Since multiple short context-response pairs could be created in one dialogue session, we eventually constructed 12M, 9M, and 6M sub-context-response pairs for Ubuntu Corpus, Douban Corpus, E-commerce Corpus, respectively. These sub-context-response pairs were used for the post-training.

4.3 Evaluation Metric

Following the previous works (Tao et al., 2019; Yuan et al., 2019; Gu et al., 2020), we used $recall$ as an evaluation metric. $Recall$ is denoted as $R_{10}@k$, which implies that the correct answer exists among the top k candidates out of the ten candidate responses. Specifically, in the experiment, $R_{10}@1$, $R_{10}@2$, and $R_{10}@5$ were used. Apart from $R_{10}@k$, we also employed MAP (mean average precision), MRR (mean reciprocal rank), and $P@1$ (precision at one) for the Douban Corpus because the dataset may contain more than one positive response from the candidates.

4.4 Baseline Methods

We compared our fine-grained post-trained model, **BERT-FP**, with the following previous models. For the initial checkpoint, we adapted the BERT$_{base}$ (110M) from Devlin et al. (2019).

- **Single-turn matching models:** Lowe et al. (2015), Kadlec et al. (2015) proposed basic models with RNN, CNN, and LSTM.

- **SMN:** Wu et al. (2017) decomposes the context-response pair into several utterance-response pairs. After matching every utterance and response, the matching vector is accumulated as the final matching score.

- **DUA:** Zhang et al. (2018) formulates the previous utterances into the context by using a deep utterance aggregation.

- **DAM:** Zhou et al. (2018) proposed a transformer encoder-based model and calculated the matching score between the context and response through self-attention and cross-attention.

- **IoI:** Through multiple interaction block chains, Tao et al. (2019) allows for deep-level matching between the utterances and responses.

- **ESIM:** Chen and Wang (2019) applied the neural language inference (NLI)'s ESIM model to the response selection.

- **MSN:** Yuan et al. (2019)'s model selects more relevant context utterances with a multi-hop selector, and it determines the degree of matching between the selected context utterances and the response.

- **BERT:** A vanilla model fine-tuned to the response selection task on the pre-trained BERT$_{base}$ without post-training.

- **RoBERTa-SS-DA:** Lu et al. (2020) proposed the speaker segmentation approach, which discriminates the different speakers and also applied dialogue augmentation.

- **BERT-DPT:** Whang et al. (2020) proposed a model that applies domain post-training (DPT). The model is post-trained with BERT's pre-training methods, MLM and NSP, and then fine-tuned to the response selection task.

- **BERT-VFT:** Whang et al. (2020) applied the efficient variable fine-tuning (VFT) method that was proposed by Houlsby et al. (2019).

- **SA-BERT:** Gu et al. (2020) incorporated speaker-aware embedding to the model; therefore, it is aware of the speaker change information.

Models	Ubuntu			Douban						E-commerce		
	$R_{10}@1$	$R_{10}@2$	$R_{10}@5$	MAP	MRR	$P@1$	$R_{10}@1$	$R_{10}@2$	$R_{10}@5$	$R_{10}@1$	$R_{10}@2$	$R_{10}@5$
TF-IDF (Lowe et al., 2015)	0.410	0.545	0.708	0.331	0.359	0.180	0.096	0.172	0.405	0.159	0.256	0.477
RNN (Lowe et al., 2015)	0.403	0.547	0.819	0.390	0.422	0.208	0.118	0.223	0.589	0.325	0.463	0.775
CNN (Kadlec et al., 2015)	0.549	0.684	0.896	0.417	0.440	0.226	0.121	0.252	0.647	0.328	0.515	0.792
LSTM (Kadlec et al., 2015)	0.638	0.784	0.949	0.485	0.537	0.320	0.187	0.343	0.720	0.365	0.536	0.828
SMN (Wu et al., 2017)	0.726	0.847	0.961	0.529	0.569	0.397	0.233	0.396	0.724	0.453	0.654	0.886
DUA (Zhang et al., 2018)	0.752	0.868	0.962	0.551	0.599	0.421	0.243	0.421	0.780	0.501	0.700	0.921
DAM(Zhou et al., 2018)	0.767	0.874	0.969	0.550	0.601	0.427	0.254	0.410	0.757	0.526	0.727	0.933
IOI (Tao et al., 2019)	0.796	0.894	0.974	0.573	0.621	0.444	0.269	0.451	0.786	0.563	0.768	0.950
ESIM (Chen and Wang, 2019)	0.796	0.894	0.975	-	-	-	-	-	-	0.570	0.767	0.948
MSN (Yuan et al., 2019)	0.800	0.899	0.978	0.587	0.632	0.470	0.295	0.452	0.788	0.606	0.770	0.937
BERT (Gu et al., 2020)	0.808	0.897	0.975	0.591	0.633	0.454	0.280	0.470	0.828	0.610	0.814	0.973
RoBERTa-SS-DA (Lu et al., 2020)	0.826	0.909	0.978	0.602	0.646	0.460	0.280	*0.495*	0.847	0.627	0.835	0.980
BERT-DPT (Whang et al., 2020)	0.851	0.924	0.984	-	-	-	-	-	-	-	-	-
BERT-VFT (Whang et al., 2020)	0.855	0.928	0.985	-	-	-	-	-	-	-	-	-
SA-BERT (Gu et al., 2020)	0.855	0.928	0.983	0.619	0.659	0.496	0.313	0.481	0.847	0.704	0.879	0.985
UMS$_{BERT+}$ (Whang et al., 2021)	0.875	0.942	0.988	*0.625*	*0.664*	*0.499*	*0.318*	0.482	*0.858*	0.762	0.905	0.986
BERT-SL (Xu et al., 2021)	*0.884*	*0.946*	*0.990*	-	-	-	-	-	-	*0.776*	*0.919*	*0.991*
BERT-FP	**0.911**	**0.962**	**0.994**	**0.644**	**0.680**	**0.512**	**0.324**	**0.542**	**0.870**	**0.870**	**0.956**	**0.993**
(diff. %p)	(+2.7)	(+1.6)	(+0.4)	(+1.9)	(+1.6)	(+1.3)	(+0.6)	(+4.7)	(+1.2)	(+9.4)	(+3.7)	(+0.2)

Table 2: Model comparison for the three benchmarks. BERT-SL and UMS$_{BERT+}$ are the previous state-of-the-art models

- **UMS$_{BERT+}$:** Whang et al. (2021) proposed a multi-task learning framework consisting of three tasks (i.e., utterance insertion, deletion, and search).

- **BERT-SL:** Xu et al. (2021) introduced four self-supervised tasks and trained the response selection model with these auxiliary tasks in a multi-task manner.

4.5 Experimental Results

Table 2 shows the performance of the proposed BERT-FP that is evaluated on three benchmarks. As you can see in the results, the proposed model outperformed all of the other models used as baselines. In comparison to the vanilla model of BERT, our model achieved an absolute improvement in $R_{10}@1$ by 10.3%p, 4.4%p, and 26%p on Ubuntu Corpus V1, Douban Corpus, and E-commerce Corpus, respectively. Compared to BERT-DPT, our model achieved an absolute improvement of 6%p in $R_{10}@1$ on the Ubuntu Corpus. These results indicate that fine-grained post-training, which reflects the dialogue's characteristics, is superior to the previous post-training. In comparison to the previous state-of-the-art models, UMS$_{BERT+}$ and BERT-SL, our model achieved an improved performance by a large margin in terms of all the metrics for the three benchmarks. These results demonstrate that our method effectively learns the semantic relevance and coherence between the internal utterances, which enhances selection performance significantly.
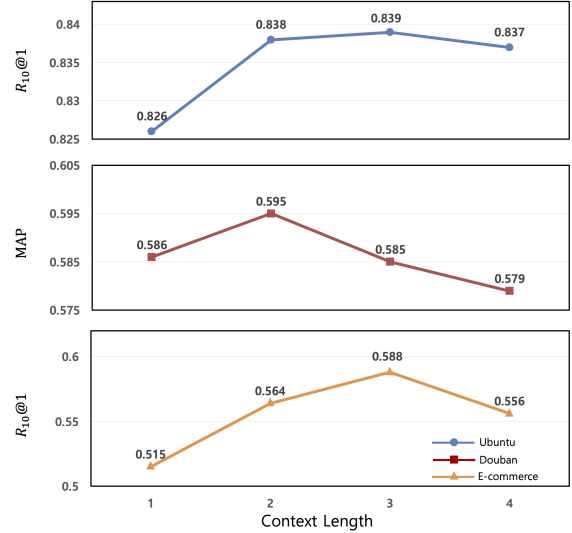

Figure 3: Performance according to post-training context length

5 Further Analysis

5.1 Performance across Different Lengths of Short Context

Figure 3 shows the performance variations of BERT-FP depending on the length of the short context. In this experiment, we trained the models with 10% of the training set and evaluated them with the entire test set to perform many experiments. Therefore, they achieved lower performance. For the Ubuntu Corpus and E-commerce Corpus, the best performance in $R_{10}@1$ is achieved when the context length is three. For Douban Corpus, we evaluated performance with MAP rather than

Models	$R_{10}@1$	$R_{10}@2$	$R_{10}@5$
BERT-NSP	0.904	0.960	0.994
BERT-SOP	0.865	0.935	0.987
BERT-URC	**0.911**	**0.962**	**0.994**

Table 3: Performance according to the training objective on Ubuntu Corpus V1.

Models	$R_{10}@1$	$R_{10}@2$	$R_{10}@5$
BERT(Gu et al., 2020)	0.808	0.897	0.975
+MLM	0.851 (+4.3%p)	0.925	0.983
+MLM+NSP	0.864 (+1.3%p)	0.934	0.987
+MLM+NSP_SCR	0.904 (+4.0%p)	0.960	0.994
+MLM+URC_SCR (Ours)	0.911 (+0.7%p)	0.962	0.994

Table 4: Ablation study on Ubuntu Corpus V1

Models	$R_{10}@1$	$R_{10}@2$	$R_{10}@5$
BERT (Gu et al., 2020)	0.808	0.897	0.975
BERT-DA	0.880	0.946	0.990
BERT-FP	0.911	0.962	0.994

Table 5: Comparing the data augmentation on Ubuntu Corpus V1

Models	$R_{10}@1$	$R_{10}@2$	$R_{10}@5$
BERT (Gu et al., 2020)	0.808	0.897	0.975
BERT-FP-NF	0.862	0.933	0.986
BERT-FP	0.911	0.962	0.994

Table 6: The effectiveness of fine-grained post-training for response selection on Ubuntu Corpus V1

$R_{10}@1$ because it may have multiple correct responses in the candidates. The best performance in MAP on Douban Corpus is achieved when the context length is set to two.

5.2 Performance according to Training Objective

We compared the proposed training objective (URC) with the previous training objectives (NSP, SOP). Table 3 demonstrates that our training objective outperforms the other training objectives. This indicates that learning both topics and the coherence between the internal utterances is important.

5.3 Ablation Study

We investigated the impact of each part of the fine-grained post-training method through a series of ablation experiments on the Ubuntu Corpus in Table 4. The model without post-training (BERT) is used as the baseline. Then, we gradually applied our methods for post-training. +MLM indicates that the model is post-trained only with the MLM. The "_SCR" suffix denotes the model that is post-trained with the short context-response pairs. The comparison between +MLM and +MLM+NSP shows that the NSP during the existing post-training has little effect on the performance. However, as shown in the comparison between +MLM and MLM+NSP_SCR, the NSP trained with a short context-response pair significantly improved the model performance. The experimental results also showed that using URC instead of NSP enhances performance.

5.4 Comparison with the Data Augmentation

The post-training method has the effect of data augmentation. However, it differs from the usual data augmentation method, which directly augments the data in the fine-tuning step. Therefore, we compared the fine-grained post-training (BERT-FP) method with the typical data augmentation (BERT-DA) on the Ubuntu Corpus. The data augmentation strategy is similar to the method used in Chen and Wang (2019). We considered each utterance as a response and its previous utterances as its context. The experimental results are shown in Table 5. BERT-FP outperforms the data augmentation model (BERT-DA) by a 3.1%p in $R_{10}@1$. The significant improvement demonstrates the effectiveness of the proposed method in comparison to the data augmentation. Our method, including post-training and fine-tuning steps, is about 2.5 times faster than BERT-DA. In particular, the post-trained model takes much less time to fine-tune than BERT-DA, making them easy to adapt to various applications.

5.5 The Effectiveness of Fine-grained Post-Training for Response Selection Task

To demonstrate the effectiveness of the fine-grained post-training method for the response selection task, we compared three different models: BERT, BERT-FP, and BERT-FP-NF (no fine-tuning). BERT-FP-NF is a model that was post-trained and evaluated without fine-tuning. As shown in Table 6, the performance of BERT-FP-NF is close to BERT-FP, which is fine-tuned. These results show that even before fine-tuning to the response selection task, our fine-grained post-training alone could measure the matching degree between the context and the response.

6 Conclusion

In this paper, we have proposed a new fine-grained post-training method that is suitable for the multi-turn dialogue. The proposed method allows the matching model to learn the semantic relevance and the coherence of the utterances in the dialogue, and it improves the model's capability to select the appropriate response. The experimental results on the three benchmark datasets demonstrate our post-training method's superiority for the response selection. From this, our model achieved a new state-of-the-art performance for all three benchmarks.

In the future, we plan to research new post-training methods that are suitable for a variety of tasks, such as question answering and dialogue generation.

Acknowledgement

References

Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2015. Neural machine translation by jointly learning to align and translate. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

Debanjan Chaudhuri, Agustinus Kristiadi, Jens Lehmann, and Asja Fischer. 2018. Improving response selection in multi-turn dialogue systems by incorporating domain knowledge. In *Proceedings of the 22nd Conference on Computational Natural Language Learning*, pages 497–507, Brussels, Belgium. Association for Computational Linguistics.

Qian Chen and Wen Wang. 2019. Sequential attention-based network for noetic end-to-end response selection. *CoRR*, abs/1901.02609.

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, Minneapolis, Minnesota. Association for Computational Linguistics.

Jia-Chen Gu, Tianda Li, Quan Liu, Zhen-Hua Ling, Zhiming Su, Si Wei, and Xiaodan Zhu. 2020. Speaker-aware BERT for multi-turn response selection in retrieval-based chatbots. In *CIKM '20: The 29th ACM International Conference on Information and Knowledge Management, Virtual Event, Ireland, October 19-23, 2020*, pages 2041–2044. ACM.

Donghoon Ham, Jeong-Gwan Lee, Youngsoo Jang, and Kee-Eung Kim. 2020. End-to-end neural pipeline for goal-oriented dialogue systems using GPT-2. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 583–592, Online. Association for Computational Linguistics.

Ehsan Hosseini-Asl, Bryan McCann, Chien-Sheng Wu, Semih Yavuz, and Richard Socher. 2020. A simple language model for task-oriented dialogue.

Neil Houlsby, Andrei Giurgiu, Stanislaw Jastrzebski, Bruna Morrone, Quentin de Laroussilhe, Andrea Gesmundo, Mona Attariyan, and Sylvain Gelly. 2019. Parameter-efficient transfer learning for NLP. In *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, volume 97 of *Proceedings of Machine Learning Research*, pages 2790–2799. PMLR.

Samuel Humeau, Kurt Shuster, Marie-Anne Lachaux, and Jason Weston. 2020. Poly-encoders: Architectures and pre-training strategies for fast and accurate multi-sentence scoring. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net.

Rudolf Kadlec, Martin Schmid, and Jan Kleindienst. 2015. Improved deep learning baselines for ubuntu corpus dialogs. *ArXiv*, abs/1510.03753.

Zhenzhong Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, and Radu Soricut. 2020. ALBERT: A lite BERT for self-supervised learning of language representations. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

Ryan Lowe, Nissan Pow, Iulian Serban, and Joelle Pineau. 2015. The ubuntu dialogue corpus: A large dataset for research in unstructured multi-turn dialogue systems. In *SIGDIAL Conference*.

Junyu Lu, Xiancong Ren, Yazhou Ren, Ao Liu, and Zenglin Xu. 2020. Improving contextual language models for response retrieval in multi-turn conversation. In *Proceedings of the 43rd International ACM*

SIGIR conference on research and development in Information Retrieval, SIGIR 2020, Virtual Event, China, July 25-30, 2020, pages 1805–1808. ACM.

Thang Luong, Hieu Pham, and Christopher D. Manning. 2015. Effective approaches to attention-based neural machine translation. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1412–1421, Lisbon, Portugal. Association for Computational Linguistics.

Chongyang Tao, Wei Wu, Can Xu, Wenpeng Hu, Dongyan Zhao, and Rui Yan. 2019. One time of interaction may not be enough: Go deep with an interaction-over-interaction network for response selection in dialogues. In *ACL*.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA*, pages 5998–6008.

Jesse Vig and Kalai Ramea. 2019. Comparison of transfer-learning approaches for response selection in multi-turn conversations. In *Workshop on DSTC7*.

Taesun Whang, Dongyub Lee, Chanhee Lee, Kisu Yang, Dongsuk Oh, and Heuiseok Lim. 2020. An effective domain adaptive post-training method for BERT in response selection. In *Interspeech 2020, 21st Annual Conference of the International Speech Communication Association, Virtual Event, Shanghai, China, 25-29 October 2020*, pages 1585–1589. ISCA.

Taesun Whang, Dongyub Lee, Dongsuk Oh, Chanhee Lee, Kijong Han, Dong-hun Lee, and Saebyeok Lee. 2021. Do response selection models really know what's next? utterance manipulation strategies for multi-turn response selection. *Proceedings of the AAAI Conference on Artificial Intelligence*.

Xianchao Wu, Ander Martínez, and Momo Klyen. 2018. Dialog generation using multi-turn reasoning neural networks. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 2049–2059, New Orleans, Louisiana. Association for Computational Linguistics.

Yu Wu, Wei Wu, Chen Xing, Ming Zhou, and Zhoujun Li. 2017. Sequential matching network: A new architecture for multi-turn response selection in retrieval-based chatbots. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 1: Long Papers*, pages 496–505. Association for Computational Linguistics.

Hu Xu, Bing Liu, Lei Shu, and Philip Yu. 2019. BERT post-training for review reading comprehension and aspect-based sentiment analysis. 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 2324–2335, Minneapolis, Minnesota. Association for Computational Linguistics.

Ruijian Xu, Chongyang Tao, Daxin Jiang, Xueliang Zhao, Dongyan Zhao, and Rui Yan. 2021. Learning an effective context-response matching model with self-supervised tasks for retrieval-based dialogues. *Proceedings of the AAAI Conference on Artificial Intelligence*.

Z. Xu, B. Liu, B. Wang, C. Sun, and X. Wang. 2017. Incorporating loose-structured knowledge into conversation modeling via recall-gate lstm. In *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 3506–3513.

Chunyuan Yuan, Wen jie Zhou, MingMing Li, Shangwen Lv, Fuqing Zhu, Jizhong Han, and Songlin Hu. 2019. Multi-hop selector network for multi-turn response selection in retrieval-based chatbots. In *EMNLP/IJCNLP*.

Zhuosheng Zhang, Jiangtong Li, Pengfei Zhu, Hai Zhao, and Gongshen Liu. 2018. Modeling multi-turn conversation with deep utterance aggregation. In *Proceedings of the 27th International Conference on Computational Linguistics, COLING 2018, Santa Fe, New Mexico, USA, August 20-26, 2018*, pages 3740–3752. Association for Computational Linguistics.

Xiangyang Zhou, Lu Li, Daxiang Dong, Yi Liu, Ying Chen, Wayne Xin Zhao, Dianhai Yu, and Hua Wu. 2018. Multi-turn response selection for chatbots with deep attention matching network. In *ACL*.