# Two-Level Supervised Contrastive Learning for Response Selection in Multi-Turn Dialogue

**Anonymous ACL submission**

## Abstract

Selecting an appropriate response from many candidates given the utterances in a multi-turn dialogue is the key problem for a retrieval-based dialogue system. Existing work formalizes the task as matching between the utterances and a candidate and uses the cross-entropy loss in learning of the model. This paper applies contrastive learning to the problem by using the supervised contrastive loss. In this way, the learned representations of positive examples and representations of negative examples can be more distantly separated in the embedding space, and the performance of matching can be enhanced. We further develop a new method for supervised contrastive learning, referred to as two-level supervised contrastive learning, and employ the method in response selection in multi-turn dialogue. Our method exploits two techniques: sentence token shuffling (STS) and sentence re-ordering (SR) for supervised contrastive learning. Experimental results on three benchmark datasets demonstrate that the proposed method significantly outperforms the contrastive learning baseline and the state-of-the-art methods for the task.

## 1 Introduction

Recent years have observed a significant surge in research on intelligent dialogue systems that can communicate with humans in natural language. In this paper, we focus on response selection in a multi-turn dialogue, in which the system selects the most appropriate response from a set of candidates given the utterances in the dialogue history.

Nowadays, pre-trained language models such as BERT (Devlin et al., 2019), RoBERTa (Liu et al., 2020) and ELECTRA (Clark et al., 2020) have become fundamental technologies of NLP. Naturally, one can formalize response selection as a matching problem and employ a pre-trained language model to conduct the task, where the input consists of a candidate of response and the utterances in the dialogue history. There is still room for improvement, however. Usually, the cross-entropy loss is used in training of the model, which is sensitive to noisy labels (Zhang and Sabuncu, 2018) and incapable of creating a large margin between positive and negative examples (Sun et al., 2016; Elsayed et al., 2018). In other words, the model learned with the cross entropy may not be strong enough to capture the features that can distinctively separate the positive examples and negative examples. Hadsell et al. (2006) originally propose exploiting contrastive learning as a way of self-supervised learning, in which positive and negative examples are automatically created and the contrastive loss is used. Empirical results have proved the effectiveness of contrastive learning in computer vision (Wu et al., 2018; Bachman et al., 2019; Tian et al., 2019; He et al., 2019; Chen et al., 2020a,b,c) and NLP (Gao et al., 2021). Contrastive learning thus can be naturally employed to address the problem described above. Particularly, one can use the method of SimCSE to perform the task of response selection. We elaborate these related work in Appendix.

We propose a new method of supervised contrastive learning, referred to as two-level supervised contrastive learning. We linearly combine the contrastive loss and the cross-entropy loss as the loss function. In addition, we use two techniques to augment the input text data at both the token level and sentence level for positive examples, namely sentence token shuffling (STS) and sentence re-ordering (SR), to create positive examples. For negative examples, we actively leverage the high quality mis-matched candidate provided by dataset. Our method uses BERT as the pre-trained language model, takes the utterances and a candidate as input, and fine-tunes the model using the loss function mentioned above. In this way, we can create a model that can effectively separate positive examples from negative examples in the embedding space.

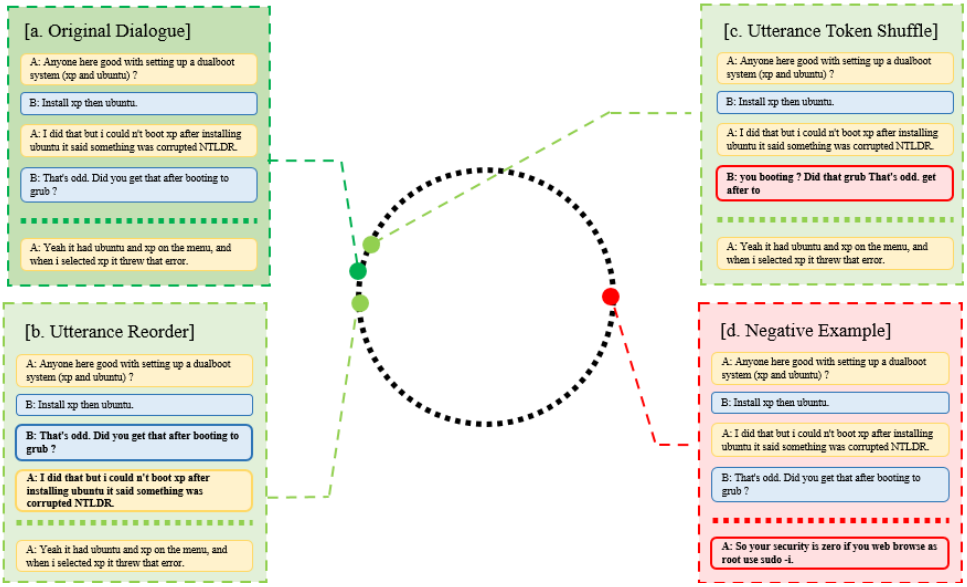


Figure 1: Example of two augmentation approaches that we integrated into contrastive learning model. The distances of hidden representations between the original dialogue and augmented example under the cosine similarity metric are expected to be closer than those between the original ones and corresponding negative examples.

We test our method on three benchmark datasets for retrieval-based dialogue: Ubuntu (Lowe et al., 2015), Douban (Wu et al., 2017), and E-Commerce Dataset (Zhang et al., 2018). Ablation study also shows the necessity of the components of our method, i.e., the loss function and the two techniques.

Our contributions are as follows.

- We propose a new supervised contrastive learning method using the contrastive loss supported by two augmentation techniques.
- We propose employing the method for enhancing the performance of response selection in multi-turn dialogue.

Experimental results show that our method outperforms existing state-of-the-art methods by a large margin.

## 2 Model and Approach

Suppose that there is a multi-turn dialogue dataset $\mathcal{D} = (c_i, r_i, y_i)_{i=1}^{N}$, where $c_i = \{u_{i,1}, u_{i,2}, u_{i,3}, \cdots, u_{i,l_i}\}$ is the context which consists of $l_i$ utterances, $r_i$ is the response for $c_i$ and $y_i \in \{0, 1\}$ is the ground truth label. We need to learn a model $F$ from $\mathcal{D}$ so that for any given context $c$ and response candidate $r$, $F(c, r) \in [0, 1]$ can measure the matching degree between $c$ and $r$.

Given a pre-trained language model $M$, we first perform domain adaptive pre-training using $\mathcal{D}$ to make the representation more domain-relevant and applicable to downstream tasks (see Appendix A.1), then we fine-tune $M$ on $\mathcal{D}$ using a contrastive fine-tuning framework.

### 2.1 Contrastive Fine-tuning Framework

We apply contrastive learning in the fine-tuning process to leverage label information. There are three major components in our framework: a data augmentation module that generates different views for input samples, an encoder $M$ that computes representations for each input text and a loss layer on top of the encoder that combines binary cross-entropy loss and contrastive loss.

### 2.2 Data Augmentation Strategies

#### 2.2.1 Sentence Token Shuffling (STS)

Instead of shuffling all the tokens in the input sequences like (Yan et al., 2021), we randomly select one utterance in the context and shuffle all tokens to form a new training example. Part c of Figure 1 illustrates how to perform sentence token shuffling.

#### 2.2.2 Sentence Reordering (SR)

In previous work (Zhang and ying Zhao, 2021), researchers let the model detect the sentence in the wrong place or predict the correct order of the input sequence. In this work, we randomly choose two utterances in the context and exchange their positions to get the augmentation. This approach enhances PrLMs' sentence level modeling ability.

To make the augmented example more reasonable, we never change the position of the last utterance in the dialogue context. Part b of Figure 1 illustrates how to perform utterance reordering augmentation.

#### 2.2.3 Two-level Augmentation Approach (TL)

We propose a technique that combines token-level and sentence-level approaches to create noisy positive samples. Specifically, we duplicate the original batch twice to get STS and SR augmentations. By incorporating two augmentation approaches into contrastive learning, the model further improves the expressiveness of the representations.

### 2.3 Supervised Contrastive Loss

For each batch of samples $X$ with size $N$, we generate augmented batch $X^+$ as we described above, and retrieve corresponding negative batch $X^-$, which contains the same context but with misleading response provided by the dataset, to get $(X, X^+, X^-)$. Then we encode these batches with PrLMs to get hidden representations $(H, H^+, H^-)$. The supervised contrastive loss can be defined as

$$\mathcal{L}_{SCL} = -\sum_{i=1}^{N} \log \frac{e^{f(h_i, h_i^+)/\tau}}{Z(H, H^+, H^-)} \quad (1)$$

and

$$Z(H, H^+, H^-) = \sum_{j=1}^{N} \left( e^{f(h_i, h_j^+)/\tau} + e^{(f(h_i, h_j^-) + \alpha \mathbb{1}_i^j)/\tau} \right) \quad (2)$$

where $h_i, h_i^+, h_i^-$ is the $i$th hidden representation of original, augmented and negative data batch, and $f$ is the similarity function as which we adopt cosine similarity in this work. $\mathbb{1}_i^j \in \{0, 1\}$ is the indicator function that equals 1 if and only if $i = j$, and $\alpha$ is the penalty factor for hard negative samples.

Since the prevailing datasets have already provided true negative pairs, it may be unreasonable to treat these high-quality negative examples equally with other in-batch negatives. To fully leverage these hard negative, we add a penalty $\alpha$ to the similarity score of $h_i$ and $h_i^-$ so that the model can significantly contrast them.

Finally, we add the supervised contrastive loss to binary cross-entropy loss with hyper-parameter $\lambda$ and get the final objective function:

$$\mathcal{L} = \mathcal{L}_{SCL} + \lambda \mathcal{L}_{CE} \quad (3)$$

## 3 Experiments

### 3.1 Datasets

We evaluate the proposed method on three public multi-turn response selection datasets, Ubuntu Dialogue Corpus V1 Xu et al. (2016), Douban Corpus (Wu et al., 2017) and E-commerce Dialogue Corpus (Zhang et al., 2018). Some statistics of these datasets are provided in Table 4 (see Appendix A.2).

### 3.2 Evaluation Metrics

Following the previous work, we use Recall$_n$@$k$ or $R_n$@$k$, $k = \{1, 2, 5\}$ as evaluation metrics, where $n$ is the number of response candidates of each context and $k$ is the top-$k$ scored candidates among $n$ candidates. Mean average precision (MAP), mean reciprocal rank (MRR), and precision at one (P@1) are also used for Douban Corpus to evaluate the model's ability to retrieve multiple responses.

### 3.3 Experimental Results

Table 1 presents the evaluation results of our model and previous methods on three datasets. As we can see that, BERT$^+_{TL}$ outperforms the present models on most metrics. Our model outperforms the prior state-of-the-art model by 5.7% in Recall$_{10}$@1 on E-Commerce Corpus, 3.1% MAP and 3.8% MRR on Douban Corpus. BERT$^+_{TL}$ also shows a superior ability to the previous fine-tuning methods like BERT-SL. Furthermore, as shown in Table 3, in comparison to the common contrastive learning method BERT$^+_{DROP}$, BERT$^+_{TL}$ achieves an absolute improvement in Recall$_{10}$@1 by 2.3% on Douban Corpus and 1% on E-Commerce Corpus. All these results demonstrate the effectiveness of our model to get better representations and retrieve the best-matched response. Readers can refer to Appendix for training details.

## 4 Analysis

Our model differs in various ways from the previous models: it adds the supervised contrastive objective to the loss function, and it integrates a two-level augmentation technique into the supervised contrastive learning model. We analyze the impact of these components on overall performance.

To evaluate the effectiveness of supervised contrastive learning loss, we compare the models with augmentations and contrastive loss (BERT$^+_{TL}$) and models with augmentations but without contrastive

| | Data / Model Name | E-Commerce | | | Douban | | | | | | Ubuntu V1 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Recall$_{10}$@1** | **Recall$_{10}$@2** | **Recall$_{10}$@5** | **MAP** | **MRR** | **P@1** | **Recall$_{10}$@1** | **Recall$_{10}$@2** | **Recall$_{10}$@5** | **Recall$_2$@1** | **Recall$_{10}$@1** | **Recall$_{10}$@2** | **Recall$_{10}$@5** |
| non PrLM-based | LSTM (Lowe et al., 2015) | 0.365 | 0.536 | 0.828 | 0.485 | 0.537 | 0.320 | 0.187 | 0.343 | 0.720 | - | 0.638 | 0.784 | 0.949 |
| | SMN (Wu et al., 2017) | 0.453 | 0.654 | 0.886 | 0.529 | 0.569 | 0.397 | 0.233 | 0.396 | 0.724 | 0.926 | 0.726 | 0.847 | 0.961 |
| | DUA (Zhang et al., 2018) | 0.501 | 0.700 | 0.921 | 0.551 | 0.599 | 0.421 | 0.243 | 0.421 | 0.780 | - | 0.752 | 0.868 | 0.962 |
| | DAM (Zhou et al., 2018) | 0.526 | 0.727 | 0.933 | 0.550 | 0.601 | 0.427 | 0.254 | 0.410 | 0.757 | 0.938 | 0.767 | 0.874 | 0.969 |
| | MRFN (Tao et al., 2019a) | - | - | - | 0.571 | 0.617 | 0.448 | 0.276 | 0.435 | 0.783 | 0.945 | 0.786 | 0.886 | 0.976 |
| | IMN (Gu et al., 2019) | 0.621 | 0.797 | 0.964 | 0.570 | 0.615 | 0.433 | 0.262 | 0.452 | 0.789 | - | 0.794 | 0.889 | 0.974 |
| | IoI (Tao et al., 2019b) | - | - | - | 0.573 | 0.621 | 0.444 | 0.269 | 0.451 | 0.786 | 0.947 | 0.796 | 0.894 | 0.974 |
| | MSN (Yuan et al., 2019) | 0.606 | 0.770 | 0.937 | 0.587 | 0.632 | 0.470 | 0.295 | 0.452 | 0.788 | - | 0.800 | 0.899 | 0.978 |
| PrLMs-based | BERT(Gu et al., 2020) | 0.610 | 0.814 | 0.973 | 0.591 | 0.633 | 0.454 | 0.280 | 0.470 | 0.828 | 0.950 | 0.808 | 0.897 | 0.975 |
| | BERT-VFT (Whang et al., 2020) | 0.717 | 0.884 | 0.986 | - | - | - | - | - | - | 0.969 | 0.867 | 0.939 | 0.987 |
| | SA-BERT (Devlin et al., 2019) | 0.704 | 0.879 | 0.985 | 0.619 | 0.659 | 0.496 | 0.313 | 0.481 | 0.847 | 0.965 | 0.855 | 0.928 | 0.983 |
| | UMS-BERT+(Whang et al., 2021) | 0.762 | 0.905 | 0.986 | 0.625 | 0.664 | 0.499 | 0.318 | 0.482 | 0.858 | - | 0.875 | 0.942 | 0.988 |
| | ELECTRA(Whang et al., 2021) | 0.607 | 0.813 | 0.960 | 0.599 | 0.643 | 0.471 | 0.287 | 0.474 | 0.831 | 0.960 | 0.845 | 0.919 | 0.979 |
| | UMS-ELECTRA+(Whang et al., 2021) | 0.707 | 0.853 | 0.974 | 0.623 | 0.663 | 0.492 | 0.307 | 0.501* | 0.851 | - | 0.875 | 0.941 | 0.988 |
| | MDFN(Liu et al., 2021) | 0.639 | 0.829 | 0.971 | 0.624 | 0.663 | 0.498 | 0.325 | 0.511 | 0.855 | 0.967 | 0.866 | 0.932 | 0.984 |
| | BERT-SL(Xu et al., 2021) | 0.776 | 0.919 | 0.991 | - | - | - | - | - | - | 0.975 | 0.884 | 0.946 | 0.990 |
| | BERT-FP(Han et al., 2021) | 0.870 | 0.956 | 0.993 | 0.644 | 0.680 | 0.512 | 0.324- | 0.542 | 0.870 | - | **0.911** | **0.962** | **0.994** |
| Our Model | BERT$^+_{TL}$ | **0.927** | **0.974** | **0.997** | **0.675** | **0.718** | **0.564** | **0.367** | **0.571** | **0.874** | **0.981** | 0.910 | **0.962** | 0.993 |
| | (diff. %p) | (5.7) | (2.3) | (0.4) | (3.1) | (3.8) | (5.2) | (4.3) | (2.9) | (0.4) | (0.6) | (-0.1) | (0.0) | (-0.1) |

Table 1: Results of previous methods and our model on E-Commerce Corpus, Douban Corpus, and Ubuntu V1 Corpus. All the results except ours are from the existing literature. The last row shows the difference between our result and prior state-of-the-art result.

| Data | Model \ Metrics | **Recall$_{10}$@1** | **Recall$_{10}$@2** | **Recall$_{10}$@5** | **P@1** | **MAP** | **MRR** |
|---|---|---|---|---|---|---|---|
| E-Commerce | BERT$^+$ | 0.903 | 0.973 | 0.998 | - | - | - |
| | BERT$^+_{TL}$ w/o CL | 0.914(+1.1%) | 0.976(+0.3%) | 0.999(+0.1%) | - | - | - |
| | BERT$^+_{TL}$ | 0.927(+2.4%) | 0.974(+0.1%) | 0.997(-0.1%) | - | - | - |
| Douban | BERT$^+$ | 0.342 | 0.558 | 0.870 | 0.535 | 0.657 | 0.700 |
| | BERT$^+_{TL}$ w/o CL | 0.351(+0.9%) | 0.562(+0.4%) | 0.878(+0.8%) | 0.546(+1.1%) | 0.664(+0.7%) | 0.707(+0.7%) |
| | BERT$^+_{TL}$ | 0.367(+2.5%) | 0.571(+1.3%) | 0.874(+0.4%) | 0.564(+2.9%) | 0.675(+1.8%) | 0.718(+1.8%) |
| Ubuntu | BERT$^+$ | 0.907 | 0.959 | 0.992 | - | - | - |
| | BERT$^+_{TL}$ w/o CL | 0.907(+0.0%) | 0.960(+0.1%) | 0.992(+0.0%) | - | - | - |
| | BERT$^+_{TL}$ | 0.910(+0.3%) | 0.962(+0.3%) | 0.993(+0.1%) | - | - | - |

Table 2: The effectiveness of different components on three datasets.

loss (BERT$^+_{TL}$ w/o CL). We use BERT$^+$, the BERT-base model after post-training and standard fine-tuning process, as the baseline. Table 2 shows that the models with two-level augmentations have already outperformed those without augmentations, and training with contrastive learning loss further boosts the performance.

| Data | Model \ Metrics | **Recall$_{10}$@1** | **Recall$_{10}$@2** | **Recall$_{10}$@5** | **P@1** | **MAP** | **MRR** |
|---|---|---|---|---|---|---|---|
| E-Commerce | BERT$^+$ | 0.903 | 0.973 | 0.998 | - | - | - |
| | BERT$^+_{DROP}$ | 0.917 | 0.974 | 0.998 | - | - | - |
| | BERT$^+_{SR}$ | 0.925 | 0.976 | 0.997 | - | - | - |
| | BERT$^+_{STS}$ | 0.919 | **0.978** | **0.999** | - | - | - |
| | BERT$^+_{TL}$ | **0.927** | 0.974 | 0.997 | - | - | - |
| Douban | BERT$^+$ | 0.342 | 0.558 | 0.870 | 0.535 | 0.657 | 0.700 |
| | BERT$^+_{DROP}$ | 0.344 | 0.562 | 0.871 | 0.545 | 0.707 | 0.664 |
| | BERT$^+_{SR}$ | 0.353 | 0.564 | **0.878** | 0.547 | 0.709 | 0.666 |
| | BERT$^+_{STS}$ | 0.355 | **0.575** | 0.870 | 0.549 | 0.708 | 0.667 |
| | BERT$^+_{TL}$ | **0.367** | 0.571 | 0.874 | **0.564** | **0.718** | **0.675** |
| Ubuntu | BERT$^+$ | 0.907 | 0.959 | 0.992 | - | - | |
| | BERT$^+_{DROP}$ | 0.909 | 0.961 | **0.993** | - | - | - |
| | BERT$^+_{SR}$ | **0.910** | **0.962** | **0.993** | - | - | - |
| | BERT$^+_{STS}$ | **0.910** | 0.961 | 0.992 | - | - | - |
| | BERT$^+_{TL}$ | **0.910** | **0.962** | **0.993** | - | - | - |

Table 3: Comparing models with different augmentation techniques on three datasets.

To investigate which augmentation technique in contrastive learning is more beneficial for response selection, we compare different data augmentation approaches, including Dropout (DROP), Sentence Token Shuffling (STS), Sentence Reordering (SR) and Two-Level augmentation (TL) in Table 3. BERT$^+$ is used as the baseline. BERT$^+_{DROP}$, BERT$^+_{STS}$ and BERT$^+_{SR}$ are the models fine-tuned using simple supervised contrastive learning loss with DROP, STS and SR as the augmentation strategy. Contrastive learning with standard dropout as noise, is a technique first proposed in SimCSE (Gao et al., 2021), and widely used in contrastive learning models in NLP. The results show that the simple method BERT$^+_{DROP}$ has already achieved consistent improvements across three datasets. For further comparison, we obtain the conclusion that STS and SR are more effective than DROP strategy and in most cases STS is slightly better than SR, indicating the representation space learned with sentence-level augmentation technique is more expressive in a multi-turn dialogue setting. Moreover, we find that the two-level augmentation technique further improves the performance by 0.9% in MAP and MRR on Douban Corpus.

## 5 Conclusion

We propose a new supervised contrastive learning method and apply it to the task of response selection in multi-turn dialogues. We show that our method outperforms existing approaches on multiple datasets: Ubuntu V1, Douban, and E-Commerce. Particularly, we significantly boost the performances of the state-of-the-art methods on Douban (5.8%) and E-Commerce (4.3%) in terms of $R10@1$.

Our method called two-level supervised contrastive learning makes use of two techniques, Sentence Token Shuffling and Sentence Reordering. We demonstrate the effectiveness of our method through an ablation study. Our method can be potentially applied to other tasks such as reading comprehension and intent classification. We plan to conduct further investigations in the future.

## References


Philip Bachman, R Devon Hjelm, and William Buchwalter. 2019. Learning representations by maximizing mutual information across views. *arXiv preprint arXiv:1906.00910*.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. 2020a. A simple framework for contrastive learning of visual representations. *arXiv preprint arXiv:2002.05709*.

Ting Chen, Simon Kornblith, Kevin Swersky, Mohammad Norouzi, and Geoffrey Hinton. 2020b. Big self-supervised models are strong semi-supervised learners. *arXiv preprint arXiv:2006.10029*.

Xinlei Chen, Haoqi Fan, Ross Girshick, and Kaiming He. 2020c. Improved baselines with momentum contrastive learning. *arXiv preprint arXiv:2003.04297*.

Kevin Clark, Minh-Thang Luong, Quoc V. Le, and Christopher D. Manning. 2020. Electra: Pre-training text encoders as discriminators rather than generators. In *International Conference on Learning Representations*.

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.

Jesse Dodge, Maarten Sap, Ana Marasovic, William Agnew, Gabriel Ilharco, Dirk Groeneveld, and Matt Gardner. 2021. Documenting the english colossal clean crawled corpus.

Gamaleldin F. Elsayed, Dilip Krishnan, Hossein Mobahi, Kevin Regan, and Samy Bengio. 2018. Large margin deep networks for classification. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 850–860, Red Hook, NY, USA. Curran Associates Inc.

Hongchao Fang and P. Xie. 2020. Cert: Contrastive self-supervised learning for language understanding. *arXiv: Computation and Language*.

Tianyu Gao, Xingcheng Yao, and Danqi Chen. 2021. SimCSE: Simple contrastive learning of sentence embeddings.

John Giorgi, Osvald Nitski, Bo Wang, and Gary Bader. 2021. DeCLUTR: Deep contrastive learning for unsupervised textual representations. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 879–895, Online. 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*, page 2041–2044. Association for Computing Machinery, New York, NY, USA.

Jia-Chen Gu, Zhen-Hua Ling, and Quan Liu. 2019. Interactive matching network for multi-turn response selection in retrieval-based chatbots. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, CIKM '19, page 2321–2324, New York, NY, USA. Association for Computing Machinery.

Beliz Gunel, Jingfei Du, Alexis Conneau, and Veselin Stoyanov. 2021. Supervised contrastive learning for pre-trained language model fine-tuning. In *International Conference on Learning Representations*.

R. Hadsell, S. Chopra, and Y. LeCun. 2006. Dimensionality reduction by learning an invariant mapping. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, volume 2, pages 1735–1742.

Janghoon Han, Taesuk Hong, Byoungjae Kim, Youngjoong Ko, and Jungyun Seo. 2021. Fine-grained post-training for improving retrieval-based dialogue systems. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1549–1558, Online. Association for Computational Linguistics.

Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick. 2019. Momentum contrast for unsupervised visual representation learning. *arXiv preprint arXiv:1911.05722*.

Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. 2020. Supervised contrastive learning. *arXiv preprint arXiv:2004.11362*.

Longxiang Liu, Zhuosheng Zhang, Hai Zhao, Xi Zhou, and Xiang Zhou. 2021. Filling the gap of utterance-aware and speaker-aware representation for multi-turn dialogue. In *AAAI*.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2020. Ro{bert}a: A robustly optimized {bert} pretraining approach.

Yixin Liu and Pengfei Liu. 2021. SimCLS: A simple framework for contrastive learning of abstractive summarization. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 1065–1072, Online. Association for Computational Linguistics.

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 *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 285–294, Prague, Czech Republic. Association for Computational Linguistics.

Lin Pan, Chung-Wei Hang, Avirup Sil, Saloni Potdar, and Mo Yu. 2021. Improved text classification via contrastive adversarial training.

Alec Radford, Jeff Wu, R. Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. Language models are unsupervised multitask learners.

Shizhao Sun, Wei Chen, Liwei Wang, Xiaoguang Liu, and Tie-Yan Liu. 2016. On the depth of deep neural networks: A theoretical view. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, AAAI'16, page 2066–2072. AAAI Press.

Chongyang Tao, Wei Wu, Can Xu, Wenpeng Hu, Dongyan Zhao, and Rui Yan. 2019a. Multi-representation fusion network for multi-turn response selection in retrieval-based chatbots. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, WSDM '19, page 267–275, New York, NY, USA. Association for Computing Machinery.

Chongyang Tao, Wei Wu, Can Xu, Wenpeng Hu, Dongyan Zhao, and Rui Yan. 2019b. One time of interaction may not be enough: Go deep with an interaction-over-interaction network for response selection in dialogues. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 1–11, Florence, Italy. Association for Computational Linguistics.

Yonglong Tian, Dilip Krishnan, and Phillip Isola. 2019. Contrastive multiview coding. *arXiv preprint arXiv:1906.05849*.

Wei Wang, Bin Bi, Ming Yan, Chen Wu, Jiangnan Xia, Zuyi Bao, Liwei Peng, and Luo Si. 2020. Structbert: Incorporating language structures into pre-training for deep language understanding. In *International Conference on Learning Representations*.

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. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2020-October:1585–1589.

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*, 35(16):14041–14049.

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 (Volume 1: Long Papers)*, pages 496–505, Vancouver, Canada. Association for Computational Linguistics.

Zhirong Wu, Yuanjun Xiong, X Yu Stella, and Dahua Lin. 2018. Unsupervised feature learning via non-parametric instance discrimination. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

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. In *AAAI*.

Zhen Xu, B. Liu, Baoxun Wang, Chengjie Sun, and Xiaolong Wang. 2016. Incorporating loose-structured knowledge into lstm with recall gate for conversation modeling. *ArXiv*, abs/1605.05110.

Yuanmeng Yan, Rumei Li, Sirui Wang, Fuzheng Zhang, Wei Wu, and Weiran Xu. 2021. Consert: A contrastive framework for self-supervised sentence representation transfer. *arXiv preprint arXiv:2105.11741*.

Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Russ R Salakhutdinov, and Quoc V Le. 2019. Xlnet: Generalized autoregressive pretraining for language understanding. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc.

Chunyuan Yuan, Wei 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 *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 111–120, Hong Kong, China. Association for Computational Linguistics.

Zhilu Zhang and Mert R. Sabuncu. 2018. Generalized cross entropy loss for training deep neural networks with noisy labels. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 8792–8802, Red Hook, NY, USA. Curran Associates Inc.

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*, pages 3740–3752, Santa Fe, New Mexico, USA. Association for Computational Linguistics.

Zhuosheng Zhang and Hai ying Zhao. 2021. Structural pre-training for dialogue comprehension. In *ACL/IJCNLP*.

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 *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1118–1127, Melbourne, Australia. Association for Computational Linguistics.

Yukun Zhu, Ryan Kiros, Rich Zemel, Ruslan Salakhutdinov, Raquel Urtasun, Antonio Torralba, and Sanja Fidler. 2015. Aligning books and movies: Towards story-like visual explanations by watching movies and reading books. In *The IEEE International Conference on Computer Vision (ICCV)*.

# A Appendix

## A.1 Domain Adaptive Post Training

Since most of pre-trained language models (PrLMs) are trained on general corpus (e.g., Wikipedia, Book Corpus (Zhu et al., 2015), Common Crawls (Dodge et al., 2021)), it is insufficient to have enough supervision of domain-specific words during fine-tuning. To help the model understand domain-specific data, we perform domain adaptive post-training using the original dialogue corpus with pre-training tasks.

Inspired by prior work (Han et al., 2021), we utilize the following procedure to generate training examples to fully exploit the training data.

1. Retrieve a random example $(c_i, r_i, y_i)$ where $y_i = 1$, and let dialogue $d_i = c_i \cup \{r_i\}$

2. With 50% possibility, randomly cut a part of $d_i$ as new context $d_i'$ which contains at least 2 utterance, for another 50% possibility, let $d_i' = d_i$

3. Sample the sentence for Next Sentence Prediction (NSP). With 25% probability, we generate a positive example with the last utterance of $d_i$ as response and the rest as context. For negative examples, 2/3 of the false responses are utterances chosen from corpus at random, and 1/3 are utterances chosen from the same context. We denote this example as $d_i''$

4. Convert $d_i''$ to training example $e_i$ in form of BERT input.

5. Perform BERT-style whole word masking on $e_i$ for MLM task

| | Part | Ubuntu V1 | Douban | E-Commerce |
|---|---|---|---|---|
| #pairs | Train | 1M | 1M | 1M |
| | Dev | 500K | 50K | 10K |
| | Test | 500K | 6670 | 10K |
| Pos:Neg | Train | 1:1 | 1:1 | 1:1 |
| | Dev | 1:9 | 1:1 | 1:1 |
| | Test | 1:9 | 1:4.62 | 1:9 |

Table 4: Data statistics for the Ubuntu, Douban, and E-Commerce Corpus.

## A.2 Datasets

### A.2.1 Ubuntu V1

Ubuntu Corpus V1 is a large multi-turn dialogue dataset extracted from Ubuntu IRC, including technical discussions of the Ubuntu system. We use the version generated by Xu et al. (2016).

### A.2.2 Douban

Douban Corpus (Wu et al., 2017) is an open-domain Chinese multi-turn dialogue dataset crawled from the popular Douban SNS. The test set of the dataset is human-labeled.

### A.2.3 E-Commerce

E-commerce Corpus (Zhang et al., 2018) is a Chinese multi-turn dialogue dataset collected from real-world customer services in Taobao, which is the largest e-commerce platform in China.

The detailed profile of these three data is shown in Table 4. In the training set of all data, one dialogue context is matched with one correct response and one incorrect response. We can use the false response given by the dataset as a hard negative result.

## A.3 Implementation Details

### A.3.1 Domain Adaptive Training

We use the English BERT-base model released by Google for the English dataset Ubuntu V1 Corpus and the Chinese BERT model from google for Chinese datasets Douban Corpus and E-Commerce Corpus. For all datasets, we use a batch size of 128, a learning rate of 3e-5 with linear decay, and a warmup ratio of 0.01. We perform domain-adaptive post-training on each dataset for 2.5 million steps.

### A.3.2 Contrastive Fine-tuning

For contrastive fine-tuning, we fine-tune each model for 3 epochs and evaluate the result every

2000 steps. For Ubuntu Corpus, we use the original batch size of 48, a learning rate of 2e-5 with linear decay, a hard negative penalty $\alpha = 1.0$, and the temperature $\tau = 1.0$. For Douban Corpus, we use the original batch size of 43, a learning rate of 1e-5 with linear decay, a hard negative penalty $\alpha = 1.0$, and the temperature $\tau = 0.05$. For E-Commerce Corpus, we use the original batch size of 43, a learning rate of 2e-5 with linear decay, a hard negative penalty $\alpha = 1.0$, and the temperature $\tau = 0.05$. For all datasets, we set $\lambda = 1.0$ to combine contrastive loss and cross entropy loss.

All codes are implemented based on Huggingface Transformer library and pytorch. We perform all experiment on 8 NVIDIA Tesla V100 and fp16 mix precision calculation is applied to all experiments.

## A.4 Related Work

### A.4.1 Pre-trained Language Models

Pretraining and fine-tuning have become a new paradigm of natural language processing and understanding. GPT (Radford et al., 2019), BERT (Devlin et al., 2019), RoBERTa (Liu et al., 2020), XLNet (Yang et al., 2019), and ELECTRA (Clark et al., 2020) are prevailing pre-trained language models. GPT was trained using uni-directional language model objective, while BERT was trained in a bidirectional way using NSP and MLM tasks. RoBERTa trained the language model with more data and removed the NSP objective from the training task. ELECTRA introduced novel generator-discriminator architecture for language models to improve training efficiency. StructBERT (Wang et al., 2020) incorporated the word structural objective and sentence structural objective to leverage language structures at the word and sentence levels.

### A.4.2 Multi-Turn Dialogue Response Selection

Before the advent of pre-trained language models, researchers used CNNs and various RNNs to process the dialogue. Early work focused on the word or sentence matching. Lowe et al. (2015) first employed LSTM and CNNs in response selection task and proposed Ubuntu V1 dataset. Later work usually encoded and matched dialogues at different levels to gain better performance. Wu et al. (2017) proposed sequential matching model and Douban Dataset, Zhang et al. (2018) proposed the deep utterance aggregation and E-Commerce dataset. After self-attention and Transformer architecture were introduced, studies began to leverage self-attention architecture such as Deep Attention Network (Zhou et al., 2018) and Multi Representation Fusion Network to perform matching. Interaction over Interaction Tao et al. (2019b) designed deep interaction module and Multi-hop Selection Network Yuan et al. (2019) focused on selecting valid context for matching.

After PrLMs are proposed, they are directly applied to response selection since dialogues can also be considered as texts. To further improve the performance of PrLMs on dialogue response selection, researchers attempt to add components to PrLMs. SA-BERT (Gu et al., 2020) added speaker embedding to the model, and MDFN (Liu et al., 2021) added more dialogue-aware response interaction module. Domain Adaptive Post Training is also an effective approach for dialogue response selection, and this approach was first proposed by Whang et al. (2020). Previous state-of-the-art model BERT-FP (Han et al., 2021) used a fine-grained post-training approach to improve the performance and Structural Pre-training for Dialogue Comprehension was used. Multi-Task Learning is also an effective way, UMS (Whang et al., 2021) and BERT-SL (Xu et al., 2021) are two examples of applying self-supervised multi-task learning to dialogue response selection.

### A.4.3 Contrastive Learning

Contrastive learning is first introduced by Hadsell et al. (2006) and re-emerges in Computer Vision (CV) and Natural Language Processing (NLP) recently. In computer vision, SimCLR (Chen et al., 2020a) proposed a simple framework that can be applied to CV and Khosla et al. (2020) proved the effectiveness of contrastive learning under supervised scenarios. In NLP, CERT (Fang and Xie, 2020) was proposed to improve language understanding, while ConSERT (Yan et al., 2021), DeCuLTR (Giorgi et al., 2021), and SimCSE (Gao et al., 2021) were proposed to give good representation of sentence. In other tasks, SimCLS (Liu and Liu, 2021) applied contrastive learning in summarization, Pan et al. (2021) performed contrastive learning in text classification with adversarial training and Gunel et al. (2021) proved that fine-tuning PrLMs can also improve the performance.