Supplementary Materials of "KnowPrompt: Knowledge-aware Prompt-tuning with Synergistic Optimization for Relation Extraction"

A Detailed Statistics of Dataset

For comprehensive experiments, we carry out our experiments on five relaction extraction datasets: TACREV (Alt, Gabryszak, and Hennig 2020), Re-TACRED (Stoica, Platanios, and Póczos 2021), SemEval 2010 Task 8 (SemEval) (Hendrickx et al. 2010), Wiki80 (Han et al. 2019) and DialogRE (Yu et al. 2020). A brief introduction to these data is as follows:

TACRED-Revisit: one dataset built based on the original TACRED dataset. They find out and correct the errors in the original development set and test set of TACRED, while the training set was left intact.

Re-TACRED: another version of TACRED dataset. They address some shortcomings of the original TACRED dataset, refactor its training set, development set and test set. Re-TACRED also modifies a few relation types, finally resulting in a dataset with 40 relation types.

SemEval: a traditional dataset in relation classification containing $10,717$ annotated examples covering 9 relations with two directions and one special relation "no_relation".

WiKi80: Wiki80 is derived from FewRel (Han et al. 2018), a large scale few-shot dataset. It contains 80 relations and 56,000 instances from Wikipedia and Wikidata. Since Wiki80 is not an official benchmark, we merely evaluate it in the low-resource setting

DialogRE: DialogRE is the first human-annotated dialogue-level RE dataset. It contains 1,788 dialogues originating from the complete transcripts of a famous American television situation comedy. It is multi-label classification, as each entity pair may posses more than one relation.

B Implementation Details

This section details the training procedures and hyperparameters for each of the datasets. We utilize Pytorch to conduct experiments with 8 Nvidia 3090 GPUs. All optimizations are performed with the AdamW optimizer with a linear warmup of learning rate over the first 10% of gradient updates to a maximum value, then linear decay over the remainder of the training. Gradients are clipped if their norm exceeded 1.0, λ and weight decay on all non-bias parameters are set to 0.001 and 0.01. A grid search is used for hyperparameter tuning (maximum values bolded below).

B.1 Standard Supervised Setting

The hyper-parameter search space is shown as follows:

- learning rate lr_1 of synergistic optimization for virtual template and anchor words. [5e-5,**1e-4**, 2e-4]
- learning rate lr_2 of optimization for overall parameters. [1e-5, 2e-5, **3e-5**, 5e-5]
- number epochs 5 (for dialogre as 20)
- batch size: 16 (for tacrev, retacred and dialogre as 8)
- max seq length: 256 (for tacrev, retacred and dialogre as 512)
- gradient accumulation steps: 1 (for dialogre as 4)

B.2 Low-Resource Setting

The hyper-parameter search space is shown as follows:

- learning rate lr_1 of synergistic optimization for virtual template and anchor words: [5e-5,**1e-4**, 2e-4]
- learning rate lr_2 of optimization for overall parameters: [1e-5, **2e-5**, 3e-5, 5e-5]
- number of epochs: 30
- batch size: 16 (for tacrev, retacred and dialogre as 8)
- max seq length: 256 (for tacrev, retacred and dialogre as 512)
- gradient accumulation steps: 1 (for dialogre as 4)

References

Alt, C.; Gabryszak, A.; and Hennig, L. 2020. TACRED Revisited: A Thorough Evaluation of the TACRED Relation Extraction Task. In *Proceedings of ACL 2020*.

Han, X.; Gao, T.; Yao, Y.; Ye, D.; Liu, Z.; and Sun, M. 2019. OpenNRE: An Open and Extensible Toolkit for Neural Relation Extraction. In *Proceedings EMNLP-IJCNLP 2019*.

Han, X.; Zhu, H.; Yu, P.; Wang, Z.; Yao, Y.; Liu, Z.; and Sun, M. 2018. FewRel: A Large-Scale Supervised Few-shot Relation Classification Dataset with State-of-the-Art Evaluation. In *Proceedings of EMNLP, 2018*.

Hendrickx, I.; Kim, S. N.; Kozareva, Z.; Nakov, P.; Ó Séaghdha, D.; Padó, S.; Pennacchiotti, M.; Romano, L.; and Szpakowicz, S. 2010. SemEval-2010 Task 8: Multi-Way Classification of Semantic Relations between Pairs of Nominals. In *Proceedings of SemEval*, 33–38.

Stoica, G.; Platanios, E. A.; and Póczos, B. 2021. Re-TACRED: Addressing Shortcomings of the TACRED Dataset. *arXiv preprint arXiv:2104.08398*.

Yu, D.; Sun, K.; Cardie, C.; and Yu, D. 2020. Dialogue-Based Relation Extraction. In *Proceedings of ACL 2020*.