Lifelong Relation Extraction for prompt engineering

Shruti Belgali and Wenrui Ma

sbelgali3@gatech.edu and wma79@gatech.edu

Abstract

Recent advances in prompt engineering have made it easier to train language models on unseen data at a fraction of the cost. However, the process to come up with templates for prompts is still a manual process. Training a model on newer data distributions or fine-tuning it to a specialized domain may need experts to create the templates. This method may turn out to be time consuming and expensive especially if the data is huge and complex. We propose a method to automatically generate such templates based on relations extracted from data. Continual learning deals with the problem of learning new tasks without forgetting the performance on previous tasks. Since we want to be able to adapt to new data, we use continual learning techniques to extract relations on it.

1 Research paper summary

1.1 Background about Continual Learning

In this setup, a model is tasked with learning multiple tasks one after the other, receiving a continuous stream of training examples. Importantly, the model doesn't have access to task labels or descriptors, meaning it doesn't know which task a given example belongs to during both training and testing phases. This mirrors real-world situations where environments change over time without providing explicit task-related information.

1.2 Summary

Current NLP models are proficient at individual tasks, but they often face challenges in adapting to new tasks while retaining knowledge of previous ones, a phenomenon referred to as lifelong learning. (Biesialska et al., 2020) and (Ke and Liu, 2023) describe cutting-edge approaches tackle this by storing previous instances in episodic memory and replaying them during training and inference. However, this method encounters three main hurdles: (1) it necessitates impractically large memory components for optimal performance, (2) it is susceptible to negative transfer, and (3) it requires numerous local adjustment steps for each test instance, causing a notable slowdown in inference speed.

This paper(Wang et al., 2020b) introduces an effective meta-lifelong framework that identifies three fundamental principles of lifelong learning techniques and integrates them in a mutually reinforcing manner. To enhance the utilization of samples, the proposed approach trains the model in a manner that enables it to acquire a superior starting point for local adaptation. Comprehensive experiments on benchmarks for tasks like text classification and question answering showcase the framework's effectiveness. It attains top-tier performance using a mere 1% of the memory capacity and reduces the performance disparity with multi-task learning. Moreover, the method effectively addresses both catastrophic forgetting and negative transfer concurrently.

2 Project description

2.1 Goals and Contributions

Our project wants to rigorously evaluate the effectiveness of various Continual Learning (CL) methods in relation to extraction. Furthermore, it conclusively demonstrates the pivotal role of Lifelong Relation Extraction in generating prompts from emerging data distributions. These prompts significantly enhance the adaptability of Language Models for downstream tasks reliant on this new data.

2.2 Methods

In this study, we employ gradient-based meta-learning methods, specifically adopting the MAML (Model-Agnostic Meta-Learning) framework introduced by (Finn et al., 2017). This approach entails the acquisition of an initial set of parameter values for the model through meta-training. At meta-test time, these initialized parameters are efficiently fine-tuned using a minimal number of gradient descent steps, enabling rapid adaptation to a novel task and facilitating the attainment of commendable performance levels. We plan to evaluate this along side (Wang et al., 2020b),(Wang et al., 2020a) and some generic continual learning methods that include memory replay methods like A-GEM (Chaudhry et al., 2018) and regularization methods like MAS (Aljundi et al., 2018).

2.3 NLP tasks

We would be handling a couple of NLP tasks such as relation extraction and Question answering. As stated above, we would be using continual learning methods for relation extraction. We would be using prompt engineering in LLMs for QA answering. We would also be using some NLP preprocessing techniques like stemming and lemmatization.

2.4 Experiments: Baseline and Evaluation

For our baseline for relationship extraction, Our comparative analysis includes the Lifelong FewRel baseline model. This benchmark serves as a reference point for evaluating the performance of our relation extraction model.

We plan to employ a comprehensive set of metrics to rigorously assess continual learning methods against each other: F1 Score: This metric provides a balanced measure of precision and recall, offering insight into the overall model performance. Precision: It gauges the accuracy of positive predictions, indicating the proportion of true positives among all predicted positives. Recall@10: This metric assesses the model's ability to retrieve relevant information, specifically focusing on the top 10 ranked results. Recall@100: Similar to Recall@10, but it extends the evaluation to the top 100 ranked results, providing a broader perspective on model performance.

Acknowledgements

References

Rahaf Aljundi, Francesca Babiloni, Mohamed Elhoseiny, Marcus Rohrbach, and Tinne Tuytelaars. 2018. Memory aware synapses: Learning what (not) to forget.

Magdalena Biesialska, Katarzyna Biesialska, and Marta R. Costa-jussa. 2020. Continual lifelong learning in natural language processing: A survey.

Arslan Chaudhry, Marc'Aurelio Ranzato, Marcus Rohrbach, and Mohamed Elhoseiny. 2018. Efficient lifelong learning with a-gem.

Chelsea Finn, Pieter Abbeel, and Sergey Levine. 2017. Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 1126–1135. PMLR.

Zixuan Ke and Bing Liu. 2023. Continual learning of natural language processing tasks: A survey.

Hong Wang, Wenhan Xiong, Mo Yu, Xiaoxiao Guo, Shiyu Chang, and William Yang Wang. 2020a. Sentence embedding alignment for lifelong relation extraction.

Zirui Wang, Sanket Vaibhav Mehta, Barnabás Póczos, and Jaime Carbonell. 2020b. Efficient meta lifelong-learning with limited memory.