

Paper Title:

Continual Learning for Natural Language Generations with Transformer Calibration

Paper Link:

<https://aclanthology.org/2022.conll-1.4/>

1 Summary**1.1 Motivation**

This paper brings out the problem of catastrophic forgetting of various online continual learning of natural language generation tasks. The writers of this paper proposed a novel transformer-based model with attention calibration to cope with the issue and enhance the performance of the model of sequential tasks. This paper aims to make a generic and lightweight feature calibration approach to solve the task-incremental continual learning problems where the models are formulated on feed-forward transformer-based function.

1.2 Contribution

This paper identifies the problem of catastrophic forgetting of online continual learning of natural language generation tasks and the models are made as feed-forward transformer-based function approximations. This model formulation of a novel task-incremental learning paradigm to train the calibrated model with an interleaved optimization process to solve the catastrophic forgetting problem.

1.3 Methodology

The first methodology of this paper is transformer calibration. In this paper, the general calibration mechanism to bring out the continual learning issue in Seq2Seq generations tasks. This calibration model adds the adaptation of the transformation function in deep transformer layers via neuron calibration. This works in both model selection and feature extraction to get rid of catastrophic forgetting in the model.

1.4 Conclusion

In this paper, the author proposed an efficient seq2seq generation model where the transformer with the calibration can dynamically adjust the function with each of the individual tasks. The authors also introduce a reproductive learning approach in addition to the iterative optimization objective which stabilizes the plasticity. In addition to that, the calibration module is very simple and does not require any task-specific characteristics. Different empirical tests give the results that this strategy performs better than the baselines. Furthermore, this paper shows that a possible gateway for future work in this field is to enhance the calibration module to multilingual pre-trained models.

2 Limitations

2.1 First Limitation

In this paper, the proposed approach and its percentage of being successful are primarily tested on specific tasks only, for example, dialog response generation and paraphrase generation. This generalization may need further research based on the different domains.

2.2 Second Limitation

In this paper, the proposed models outperformed the method over baselines. However, a more comprehensive comparative analysis and a wider range of state-of-the-art continual learning algorithms and natural language generation models could provide a more thorough understanding of its strengths and limitations.

3 Synthesis

In this paper, the author proposes a novel approach to address the task of continual learning in the creation of natural language generative models. In order to solve the catastrophic forgetting of model parameters and to simplify knowledge in different tasks the proposed method uses a general calibration mechanism to cope with the adaptive transformer functions.