

XINYU ZHAO

xinyuzhao@stu.pku.edu.cn

Education

Peking University

MS in Computer Technology - GPA: 3.52/4

Sep. 2021 – Jun. 2024

Beijing, China

Huangzhong University of Science and Technology

BA in Translation and Interpreting - GPA: 3.99/4

Sep. 2017 – Jun. 2021

Wuhan, China

Huangzhong University of Science and Technology

Minor in Computer Science and Technology

Sep. 2018 – Jun. 2020

Wuhan, China

Publications

Enhancing Dynamic Image Advertising with Vision-Language Pretraining

SIGIR 2023

Co-first Author

- Query-image pairing in multimodal sponsored search consists of image retrieval and relevance modeling. Previous separate optimizing method leads to suboptimal result, we unify them by Vision-Language Pretraining.
- Base model:** train a CLIP-based model with 1B multi-domain image-text pairs to learn general representation, outperforming SOTA in Recall@10 by 6.9% on business datasets.
- Relevance model:** further train the base model on advertising data, achieving a 4.62% increase in AUC compared to separately trained model.
- Retrieval model:** unify retrieval and relevance modeling with multitask training, improve Recall@10 and Relscore@10 by 19% and 1% compared to the base model.
- Our approach has been implemented in Baidu Search Ads, obtaining 1.04% and 1.87% increase in CPM and CTR.

Sparse MoE with Language-Guided Routing for Multilingual Machine Translation

ICLR 2024 (UNDER REVIEW)

First Author

- Current Sparse Mixture-of-Experts (SMoE) solutions neglect the language features in multilingual machine translation. Therefore, we propose **Lingual-SMoE** with language-guided routing based on language hierarchy and difficulty.
- Language hierarchy:** incorporate linguistic knowledge in routing to allocate experts into different group.
- Language difficulty:** adjust the number of activated experts for each target language in an automatic manner, according to their data resource difficulties, mitigating the potential over-/under-fitting issues.
- Sufficient experiments are conducted with various network architectures, validating the superior performance of our proposals. For instance, **Lingual-SMoE** outperforms its dense counterpart by over 5% BLEU scores on OPUS-100.

Research Experiences

Internship: Baidu Search Ads | Multimodal Deep Learning

Nov. 2022 – Jun. 2023

- Scaling the vision-language model to 1B parameter is computationally expensive, we tackle the problem from training and data aspects.
- Training efficiency:** partition optimizer states and gradients using ZeRO with Deepspeed, and randomly drop image patches during training.
- Data quality:** generate additional texts with image captioning model, then use our two-tower model to filter low-quality image-text pairs.

Course Project: Chinese Grammatical Error Correction

Jun. 2022 – Jul. 2022

- In Chinese grammatical error correction, there are many forms of grammatical errors. As relevant data is limited, it is difficult to directly apply existing pre-trained language model.
- Data augmentation:** inject noise according to data distribution and metrics to alleviate data scarcity.
- Two-stage training:** first train a model on augmented training data, then fine-tune it with validation data to improve generalization performance on inconsistent data.
- Post-processing:** conduct post-processing on predictions based on word-level editing distance, to filter out bad cases in correction. The final result exceeds the benchmark by 14 points.

Skills

Programming Languages: Python, Shell

Programming Framework: Pytorch, Huggingface Transformers, Fairseq

Languages: TOEFL (102)

Awards

Outstanding Graduate of Huangzhong University of Science and Technology

Jun. 2021

National Scholarship

Oct. 2019