Xiaofeng Zhang

TEL: 008618810687955
E-mail: xiaofng.zhang@gmail.com
Personal Page: https://zhang-guigui.github.io

BIO

I am a master's student (2020.09 –) at the Ministry of Education (MOE) Engineering Research Center of Advanced Computer Application Technology, Beihang University (China), and a generalist engineer candidate at Centrale Pékin, Groupe des Écoles Centrale (France), supervised by Prof. Zhang Xiong and Prof. Wenge Rong. I obtained my BSc degree in Information and Computational Science (*Summa Cum Laude*) from Beihang University and Centrale Pékin (2016.09 – 2020.07).

My research interest mainly relies on sparsely activated neural networks and compositional generalization. My research goal is to conceive modularized neural networks and help to measure compositional generalization. With these neural networks, we could build explainable and reliable AI systems. I also invest in the knowledge representation in large-scale pre-trained language models. I currently focus on Natural Language Processing, but these issues are prevalent in other domains.

EDUCATION

Centrale Pékin, Beihang University – Electronic & Information, MSc 2020.9 – 2023.6
Research Direction: Natural Language Processing, supervised by Prof. Zhang Xiong and Prof. Wenge Rong
Centrale Pékin, Beihang University – Information and Computational Science, BSc 2016.9 – 2020.6
GPA: 3.871/4.00, Average Score: 92.1/100, Ranking: 1/86

AWARDS

National Scholarship for Graduate Students Award (Top 1%),
Two times of National Scholarship for Undergraduate Students Award (Top 1%),
Excellent Undergraduate Student of Beijing (Top 3%), Merit Student of Beijing (Top 1%),
Shen Yuan Honor Award (10 people in Beihang University),
First prize in 32th Fengru Cup Innovation Competition.

PUBLICATIONS

Accepted:

[1] **Xiaofeng Zhang**, Yikang Shen, Zeyu Huang, Jie Zhou, Wenge Rong, Zhang Xiong. Mixture of Attention Heads: Selecting Attention Heads Per Token. In: Proceedings of EMNLP 2022

[2] Haoxing Zhang, **Xiaofeng Zhang**, Haibo Huang, Lei Yu. Prompt-Based Meta-Learning for Few-shot Text Classification. In: Proceedings of EMNLP 2022

[3] Zeyu Huang, Wenge Rong, **Xiaofeng Zhang**, Yuanxin Ouyang, Chenghua Lin, Zhang Xiong. Token Relation Aware Chinese Named Entity Recognition. ACM Transactions on Asian and Low-Resource Language Information Processing

[4] Zhendong Wang, **Xiaofeng Zhang**, Wei Chen, Jianwei Niu. Lung Segmentation Reconstruction Based Data Augmentation Approach for Abnormal Chest X-Ray Images Diagnosis. In: Proceedings of 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) 2022

Under Review:

[5] **Xiaofeng Zhang**, Chuantao Yin, Wenge Rong, Chenghua Lin, Zhang Xiong. Subgraph-Oriented Heterogeneous Drug-Target Interaction Identification. Submitted to ACM Transactions on Computing for Healthcare, major revision

[6] Zeyu Huang, Yikang Shen, **Xiaofeng Zhang**, Jie Zhou, Wenge Rong, Zhang Xiong. Transformer-Patcher: One Mistake worth One Neuron. Submitted to ICLR 2023

INTERNSHIPS

AI Lab, Wechat, Tencent, supervised by Dr. Yikang Shen 2021.12 – 2022.6

Worked on two research topics. Firstly, we propose a new attention mechanism called Mixture of Attention Heads (MoA), which is a combination of Mixture of Experts and Multi-head Attention Mechanism. MoA permits easily scale up the model capacity while constraining the same level of computational cost. We achieved better performance on Machine Translation and Language Modeling tasks compared to SOTA baselines. Secondly, we work on the model editing task in order to efficiently correct the false prediction of a pre-trained language model. We propose a new task named sequential editing task and conceive a novel way to edit the stored knowledge in the pre-trained language model. We add a few new neurons in the last Feed Forward layer of Transformer architecture. The experimental result show that our method outperforms other strong baselines

Technology Research Centre, PingAn Technology, supervised by Dr. Yanmeng Wang 2021.7 – 2021.11

Worked on two research topics. Firstly, we propose a variant GAN to control the quality and the diversity of the generated sequences. We add a control coefficient in the GAN loss to deliberately control either the quality or the diversity. Experimental results show that our proposed GAN could converge to the performance of MLE model. Secondly, we worked on a project for summarization of the text from customer service. We basically try several SOTA baselines and adopt different DL platforms, e.g., PaddlePaddle, Huggingface.

DeeCamp, Sinovation Ventures, supervised by Dr. Ran Zhang 2021.6 – 2021.8

Worked on a multi-discipline project of Natural Language Processing and Computer-Human Interaction. Our project named BrickPal: Brick assembly with automated generated assembly sequences and AR headset. We finished a prototype that could help users to assemble brick models using an AR headset. I worked mainly on the automated sequence generation part. I used a Language Model to learn the distribution of brick assembly sequences. The results show that the automated generated assembly sequences are of high-quality and more flexible compared to traditional assembly sequences.

Beijing Lab, Orange (France), supervised by Dr. Li Jiang 2020.1 – 2020.3

Worked on a news recommendation project. We used BERT and Multilingual BERT to get the embedding of news titles. Then we computed the cosine similarity of each news title. We could recommend news for users according to their reading history.

TECHNICS AND LINGUISTICS

English: CET-4: 649/710 (TOP 1%); CET-6: 622/710 (Full band in Listening); IELTS: 8.0 (C1)
French: DELF B2: 79.5/100 (Professional Proficiency), Second Place in Trophée des talents 2020
Coding: Python, Matlab, C, Java, Latex
DL: Pytorch, PaddlePaddle, Tensorflow

INTERNATIONAL EXCHANGES

Ecole Centrale de Casablanca, Morocco – 15 days 2019.11
Tokyo Institute of Technology, Shibaura Institute of Technology, University of Tokyo, Japan – 15 days 2019.7
CentraleSupélec (The University of Paris-Saclay), France – 8 days 2018.2