Xiaoyue Xu

@ xiaoyue.xu.me@gmail.com

EDUCATION

Tsinghua University   Beijing, China
B.Sc. in Computer Science and Technology; GPA: 3.8/4.0   *Sep 2020 – Present*

PUBLICATIONS

* denotes equal contribution

1. **Boosting Inference Efficiency: Unleashing the Power of Parameter-Shared Pre-trained Language Model.** *EMNLP findings, 2023* [pdf].
   Weize Chen*, **Xiaoyue Xu***, Xu Han, Yankai Lin, Ruobing Xie, Zhiyuan Liu, Maosong Sun, Jie Zhou.

2. **Learning Heterogeneous Mixture of Hash Experts for Highly Scalable Neural Radiance Fields.** *Technical report.*
   Zhenxing Mi, **Xiaoyue Xu**, Dan Xu.

RESEARCH EXPERIENCE

Undergraduate Research Intern, THUNLP, THU
Advisor: Prof. Zhiyuan Liu   *Sep 2022 – Present*

▷ **Efficient Inference for Parameter-sharing PLMs**
- Developed a straightforward technique to significantly improve inference efficiency in parameter-sharing PLMs by utilizing an ODE perspective.
- Proposed a novel pre-training strategy, which further expedited the inference process of models with fully or partially shared parameters, retaining 99% performance at around 1.5x acceleration. (Published to EMNLP 2023 findings.)

▷ **Addressing Long-tail Distribution Problem via Hypernetwork**
- Developed a hypernetwork-based solution to bridge the gap between low-resource and high-resource scenarios, targeting long-tail data distribution challenges in NLP tasks such as relation extraction. Engineered a novel approach by modeling the training process as a SDE, generating parameter trajectories for optimal few-shot learning performance.

Visiting Research Intern, HKUST
Advisor: Prof. Dan Xu   *Jul 2023 – Sep 2023*

▷ **Transferable Monocular Depth Estimation**
- Combined relative depth pretraining and metric depth finetuning to promote model generalization capability across varied environmental conditions. Experimented on incorporating prompting methods to improve zero-shot performance.

▷ **Heterogenerous MoE for Scalable Large Scale NeRF**
- Contributed to the design and writing of a research paper which proposed a scalable and efficient large-scale NeRF framework by employing heterogeneous models with mixture of experts method.

AWARDS&HONORS

Academic Excellence Scholarship, Department of Computer Science, Tsinghua University (2021)

SKILLS

Language: Chinese(native), English(fluent, TOEFL 106)
Technology: C/C++, Python, PyTorch, Linux