

EDUCATION

School of Informatics, The University of Edinburgh	Edinburgh, UK
PhD of NLP, supervisors: Prof. Ivan Titov and Dr. Edoardo Maria Ponti	June 2027 expected
Ecole Centrale Pékin (ECPK), Beihang University	Beijing, China
Master of Electronic Information; GPA: 3.9/4.0; Ranking: 2/86	June 2023
Ecole Centrale Pékin (ECPK), Beihang University	Beijing, China
Bachelor of Information and Computing Science; GPA: 3.7/4.0; Ranking: 6/87	June 2020

SELECTED PUBLICATIONS¹

- Zihan Qiu*, **Zeyu Huang***, Bo Zheng, Kaiyue Wen, Zekun Wang, ... & Junyang Lin. Demons in the Detail: On Implementing Load Balancing Loss for Training Specialized Mixture-of-Expert Model - **ACL 2025**
- Zeyu Huang**, Zihan Qiu, Zili Wang, Edoardo M Ponti, Ivan Titov. Post-hoc Reward Calibration: A Case Study on Length Bias - **ICLR 2025**
- Zihan Qiu*, **Zeyu Huang***, Shuang Cheng, Yizhi Zhou, Zili Wang, Ivan Titov, Jie Fu. Layerwise Recurrent Router for Mixture-of-Experts - **ICLR 2025**
- Ka Man Lo*, **Zeyu Huang***, Zihan Qiu*, Zili Wang, Jie Fu. A Closer Look into Mixture-of-Experts in Large Language Models. **NAACL 2025 Findings**
- Wenyu Du, Tongxu Luo, Zihan Qiu, **Zeyu Huang**, Yikang Shen, Reynold Cheng, Yike Guo, Jie Fu. Stacking Your Transformers: A Closer Look at Model Growth for Efficient LLM Pre-Training - **NeurIPS 2024 spotlight**
- Wenyu Du, Shuang Cheng, Tongxu Luo, Zihan Qiu, **Zeyu Huang**, Ka Chun Cheung, Reynold Cheng, Jie Fu. Unlocking Continual Learning Abilities in Language Models - **EMNLP 2024 Findings**
- Zihan Qiu*, **Zeyu Huang***, Jie Fu. Unlocking Emergent Modularity in Large Language Models. **NAACL 2024. Outstanding Paper Award**
- Zeyu Huang**, Yikang Shen, Xiaofeng Zhang, Jie Zhou, Wenge Rong, Zhang Xiong. Transformer-Patcher: One Mistake Worth One Neuron. **ICLR 2023**.
- Xiaofeng Zhang*, Yikang Shen*, **Zeyu Huang**, Jie Zhou, Wenge Rong, Zhang Xiong. Mixture of Attention Heads: Selecting Attention Heads Per Token. **EMNLP 2022**

SELECTED PRE-PRINTS

- Zeyu Huang**, Tianhao Cheng, Zihan Qiu, Zili Wang, Yinghui Xu, Edoardo M Ponti, Ivan Titov. Blending Supervised and Reinforcement Fine-Tuning with Prefix Sampling - arXiv:2507.01679
- Yijun Yang*, **Zeyu Huang***, Wenhao Zhu, Zihan Qiu, Fei Yuan, Jeff Z Pan, Ivan Titov. A Controllable Examination for Long-Context Language Models - arXiv:2506.02921.
- Zihan Qiu*, Zekun Wang*, Bo Zheng*, **Zeyu Huang***, Kaiyue Wen, Songlin Yang, ... & Junyang Lin. Gated Attention for Large Language Models: Non-linearity, Sparsity, and Attention-Sink-Free. - arXiv:2505.06708.

EXPERIENCE

Student Researcher at Google Deepmind	Aug. 2025 - Dec. 2025
Advisor: Dr. Marc'Aurelio Ranzato	London
Research intern at INF Technology	Mar. 2024 - Aug. 2024
Advisor: Zili Wang	Remote
Research Intern at BAAI	Sep. 2022 - Jun. 2023
Advisor: Dr. Jie Fu	Beijing, China
Research Intern at WeChat AI	Nov. 2021 - May. 2022
Advisor: Dr. Yikang Shen	Beijing, China

HONORS AND AWARDS

- NAACL 2024 Outstanding Paper Award (6/2604 submissions)
- ESPRC DTA Scholarship (international flexibility) - 2024-2027
- National Scholarship for Graduate Students Award - 2021-2022 - (3/291)
- Outstanding Graduate Student of Beihang University - 2020-2021

¹* denotes equal contribution

- Excellent Student Cadre of Beihang University, 2020-2021
- Academic Scholarship of Beihang University, Top 10% on GPA, 2016-2019, 2020-2022
- Merit Student of Beihang University, 2017-2018, 2020-2021
- First Prize in National Olympiad in Informatics in Provinces (**NOIP**) - 2014, 2015

EXTRA-CURRICULAR ACTIVITIES

International Exchange	Casablanca, Morocco
Ecole Centrale Casablanca	Nov 2019, 15 days

International Exchange	Paris, France
CentraleSupélec, Saclay	Feb 2018, 8 days

Student works	Beijing, China
President of the Graduate Student Union of ECPK	Dec. 2020 - Dec. 2021

Orange Lab & Ecole Centrale de Pekin	Beijing, China
Oct 2020 - May 2021	Group Leader of the Project

- **Federated Recommendation** - A recommendation system trained in Federated Learning manner.
- This project is co-supervised by Orange Lab of Beijing and Ecole Centrale de Pekin.
- The product of this project includes a survey report and deployment of a Federated news recommender

SKILLS SUMMARY

- **Programming:** Python, C, Shell, JAVA, Latex, MATLAB
- **Languages:** Chinese(Native), English (IELTS 7.5), French (B2)
- **ML/AI Frameworks:** PyTorch, JAX, Tensorflow, PyTorchLightning, Fairseq, Huggingface, Flax

RELEVANT COURSES

- **Computer Science:** Programming(96), Deep Learning(1st in the school), Fundamentals of Computer Science(93)
- **Math:** Analysis(100), Probability and Statistics(99), Differential(96), Algebra(96),Mathematical Statistics(98)
- **Physics:** Structural Mechanics(100), Optics(96), Wave Physics(95), Electronics(94), Electromagnetic Radiation(95)