Qiushi Sun

MASTERS STUDENT · DATA SCIENCE

Shanghai, China / Singapore

+65 93795970 / +86 15301824319 | ✉ thomasqiushi@icloud.com | 🏠 qiushisun.github.io

Education

National University of Singapore   *Singapore*
MSc IN DATA SCIENCE   *July 2022 - January 2024*
- Natural Language Processing Research

East China Normal University   *Shanghai, China*
BENG IN DATA SCIENCE AND TECHNOLOGY   *GPA: 3.77/4.0*   *September 2018 - June 2022*
- Computer System, Data Science, Machine Learning, Distributed Systems, Database, *etc.*
- Thesis: Knowledge Distillation based on Code Pre-Trained models (Best Thesis Award)

Research Experience

Shanghai AI Laboratory   *Shanghai, China*
ADVISOR: DR. ZHIYONG WU   *June 2023 - Present*
- Research Intern: Natural Langauge Processing

Institute for Infocomm Research (I^{2}R), A*STAR   *Singapore*
ADVISOR: DR. XIAOLI LI (DEPT. HEAD & PRINCIPAL SCIENTIST)   *January 2023 - June 2023*
- Research Intern: AI for Text Data Analytics

School of Data Science and Engineering, East China Normal University   *Shanghai, China*
CO-ADVISORS: PROF. XIANG LI AND PROF. XUESONG LU   *October 2021 - October 2022*
- Lightweight Deployment / Interpretability of Code Representation Models
- Parameter-Efficient Learning for NLP

Honors and Awards

2022	**International Algorithm Case Competition: PLM Tuning, Second Prize** , Greater Bay Area	*RMB 100,000*
	Shanghai Outstanding Graduates, Shanghai Municipal Education Commission	
	Excellent Bachelor Thesis Award, East China Normal University	
2021	**Outstanding Student**, East China Normal University	
	First Class Scholarship, East China Normal University	*RMB 5000*
	Finalist (Special Class Prize), Mathematical and Interdisciplinary Contest in Modeling	
2020	**Third Class Scholarship**, East China Normal University	*RMB 1,000*
	Honarable Mention, Mathematical and Interdisciplinary Contest in Modeling	

Publications

PUBLISHED PAPERS & PREPRINTS

(* indicate equal contribution, part of the manuscripts are currently available upon request due to anonymity requirement)

[1] Nuo Chen[*], **Qiushi Sun**[*], Renyu Zhu[*], Xiang Li, Xuesong Lu and Ming Gao. *CAT-probing: A Metric-based Approach to Interpret How Pre-trained Models for PL Attend Code Structure*. EMNLP 2022, Abu Dhabi, UAE. [Paper] [Video] [Slides]

[2] Zhangyue Yin, **Qiushi Sun**, Qipeng Guo, Jiawen Wu, Xipeng Qiu and Xuanjing Huang. *Do Large Language Models Know What They Don't Know?* ACL 2023, Toronto, Canada.

[3] Chengcheng Han, Liqing Cui, Renyu Zhu, Jianing Wang, Nuo Chen, **Qiushi Sun**, Xiang Li and Ming Gao. *When Gradient Descent Meets Derivative-Free Optimization: A Match Made in Black-Box Scenario*. ACL 2023, Toronto, Canada.

[4] **Qiushi Sun**, Nuo Chen, Jianing Wang, Xiang Li and Ming Gao. *TransCoder: Towards Unified Transferable Code Representation Learning Inspired by Human Skills*. Under-Review of TACL.

[5] **Qiushi Sun**, Chengcheng Han, Nuo Chen, Renyu Zhu, Jingyang Gong, Xiang Li and Ming Gao. *Make Prompt-based Black-Box Tuning Colorful: Boosting Model Generalization from Three Orthogonal Perspectives*. Preprint.

[6] Nuo Chen, **Qiushi Sun**, Jianing Wang, Xiang Li and Ming Gao. *Pass-Tuning: Towards Structure-Aware Parameter-Efficient Tuning for Code Representation Learning*. Under-Review of TACL.

[7] Jianing Wang, **Qiushi Sun**, Nuo Chen, Chengyu Wang, Ming Gao and Jun Huang. *Uncertainty-aware Parameter-Efficient Self-training for Semi-supervised Language Understanding*. Preprint.

[8] Jianing Wang, Nuo Chen, **Qiushi Sun**, Wenkang Huang, Chengyu Wang and Ming Gao. *HugNLP: A Unified and Comprehensive Library for Natural Language Processing.* Under-Review of ACL 2023 System Demonstration. [Paper]

In Prep

Qiushi Sun, Nuo Chen, Jianing Wang and Xiaoli Li. *Probe, Attack, and Refinement: An Empirical Study of Code Pre-trained Model Robustness and Interpretability.*

Teaching Experience

Fall 2021   **Teaching Assistant**, Deep Learning for Computer Vision
Instructed by Prof. Xuesong Lu, School of Data Science and Engineering, ECNU

Projects

2020 - 2021   **"DaSE Quora"**, Portfolio
Supported by Shanghai municipal college student innovation and entrepreneurship project, Design and implement a learning guidance system to assist university students in exploring knowledge points relevant to Prob. theory and DS courses. Utilizing a KG constructed by aggregating knowledge in textbooks and Wikipedia.

Outreach & Professional Development & Skills

Outreach

Winter 2021   **Imperial College London**, Winter Exchange
Data Science & Computer Vision, IC Data Science Winter School *GPA: 4.0/4.0*

Skills

Programming/Frameworks/Skills: Python, PyTorch, Keras, SQL LaTeX, Flask, Huggingface, Technical Writing

Languages: Mandarin & Shanghainese (Native), English (Fluent, TOEFL-108, GRE-334), German (Basic)

Research Interest

My research interests cover several aspects of natural language processing, with my current focus on:

- Efficient Methods for NLP.
- Large Language Models for Code.
- Chain of Thought Reasoning for Large Language Models.

Professional Service

Reviewer for EMNLP 2022, ACL 2023