

Tianjian Li

Updated April 7, 2024

Email: tli104@jhu.edu Website: tianjianl.github.io

Education

Johns Hopkins University Baltimore, MD
PhD in Computer Science 2024 – Present
Advisor: Daniel Khashabi

Johns Hopkins University Baltimore, MD
MSE in Computer Science 2022 – 2024
Advisors: Kenton Murray, Daniel Khashabi, Philipp Koehn

New York University New York, NY
BA in Computer Science and Mathematics 2017 – 2021

Publications

Error Norm Truncation: Robust Training in the Presence of Data Noise for Text Generation Models

Tianjian Li, Haoran Xu, Philipp Koehn, Daniel Khashabi, Kenton Murray
*International Conference on Learning Representations (ICLR), 2024. **Spotlight.***

Why Does Zero-Shot Cross-Lingual Generation Fail? An Explanation and a Solution

Tianjian Li, Kenton Murray
Association of Computational Linguistics (ACL) - Findings, 2023.

Research experience

Johns Hopkins University 2022 – 2024
Center for Language and Speech Processing (CLSP)
Advisors: Kenton Murray, Daniel Khashabi, Philipp Koehn
- Efficient online domain weighting for training language models.
- (ICLR '24) Removing data errors during training for language models.
- (ACL '23) Improving cross-lingual transfer for text generation.

Tsinghua University Spring 2022
Advisor: Jie Tang
- Pre-training of large multilingual language models.
- Cross-lingual summarization with multilingual LMs.

Teaching experience

Course Assistant, Department of Computer Science Spring 2024
CSCI 601.471/671 NLP: Self-Supervised Models

Industry experience

Baidu Inc. Baidu Maps Beijing, China
Machine Learning Engineer (Intern) Fall 2021
- Optimization of estimated trip time with Graph Neural Networks.

Skills

Programming: Python, C/C++, Java
Frameworks: PyTorch, Huggingface (Accelerate), Fairseq, DeepSpeed, Jax
Tools: Docker, git, Hadoop streaming, Spark, Vim, LaTeX
Languages: Chinese (Native), English, French (Beginner)

Service

Reviewer: ACL (2023, 2024), EMNLP (2023), EACL (2024), COLM (2024)
Organizer: MASC-SLL (2024)