Giulio Zhou

Email: giuliozhou8@gmail.com | **Website:** giuliozhou.com | **GitHub**: github.com/giulio-zhou

Education

Carnegie Mellon University (Pittsburgh, PA) 08/2017 – *present*
PhD, Computer Science, *Advisor: Dave Andersen*

University of California, Berkeley (Berkeley, CA) 08/2012 – 12/2016
Bachelor of Arts, Computer Science **GPA: 3.893 / 4.0**

Relevant Coursework: Machine Learning, Artificial Intelligence, Convex Optimization, Computer Vision, Deep Reinforcement Learning, Operating Systems, Computer Networking, Database Systems.

Publications

Transforming Sentences in Latent Space with Transformer Autoencoders
Giulio Zhou, Byron Wallace, Zachary Lipton. Under Submission.

A Field Test of Bandit Algorithms for Recommendations: Understanding the Validity of Assumptions on Human Preferences in Multi-armed Bandits
Liu Leqi*, ***Giulio Zhou***, Fatma Kilinc Karzan, Zachary Lipton, Alan Montgomery. Under Submission.

Learning on Distributed Traces for Datacenter Storage Systems
Giulio Zhou, Martin Maas. Conference on Machine Learning and Systems (MLSys), 2021.

Multi-Vector Attention Models for Deep Re-ranking
Giulio Zhou, Jacob Devlin. Empirical Methods in Natural Language Processing (EMNLP), 2021.

Multi-Task Learning for Storage Systems
Giulio Zhou, Martin Maas. Workshop on ML for Systems at NeurIPS, 2019.

Scaling Video Analytics on Constrained Edge Nodes
Christopher Canel, Thomas Kim, ***Giulio Zhou***, Conglong Li, Hyeontaek Lim, David G. Andersen, Michael Kaminsky, Subramanya R. Dulloor.
Conference on Systems and Machine Learning (SysML), 2019.

Clipper: A Low-Latency Online Prediction Serving System
Daniel Crankshaw, Xin Wang, ***Giulio Zhou***, Michael J. Franklin, Joseph E. Gonzalez, Ion Stoica.
USENIX Symposium on Networked Systems Design and Implementation (NSDI), 2017.

Industry Experience

Google Inc., NLP Team (Research Intern, Host: Jacob Devlin) 06/2020 – 09/2020
- Developed efficient models for text retrieval using lightweight attention mechanisms for query-document scoring.

Google Inc., Brain Team (Research Intern, Host: Martin Maas) 05/2019 – 08/2019
- Devised a method for learning storage policies from distributed traces using Transformers.
- Demonstrated end-to-end SSD caching and SSD/HDD file placement improvements in simulation.

Google Inc., Keyboard Team (Software Engineer) 03/2017 – 08/2017
- Created infrastructure to support analytics data pipelines and interactive data visualization.
- Automated the generation, management and evaluation of Google Keyboard test sets.

Google Inc., Display Ad Automation Team (Software Engineering Intern) 05/2015 – 08/2015
- Built a backend pipeline for automated text-to-image keyword matching for internationalized display ads.

Teaching Experience

15-849, Machine Learning Systems Seminar (Spring 2022)

10-701, Introduction to Machine Learning (Spring 2020)

CS 189/289A, Introduction to Machine Learning (Fall 2016)

CS 61BL, Data Structures and Programming Methodology (Summer 2016)

CS 61B, Data Structures and Algorithms (Spring 2016)