

# 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.

## 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 Systems Team (Research Intern, Host: Martin Maas) 05/2019 – 06/2020

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.

## Publications

---

### Conference Publications

#### 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. [Short paper]

#### 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 Machine Learning and Systems (MLSys), 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.

### Under Submission

#### Transforming Sentences in Latent Space with Transformer Autoencoders

Giulio Zhou, Byron Wallace, Zachary Lipton.

#### 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. [\*Equal contribution]

### Workshop Publications

#### Multi-Task Learning for Storage Systems

Giulio Zhou, Martin Maas. Workshop on ML for Systems at NeurIPS, 2019.

#### Efficient Multi-Tenant Inference on Video using Microclassifiers

Giulio Zhou, Thomas Kim, Christopher Canel, Conglong Li, Hyeontaek Lim, David G. Andersen, Michael Kaminsky, Subramanya R. Dulloor. Conference on Machine Learning and Systems (MLSys), 2018.

#### Picking interesting frames in streaming video

Christopher Canel, Thomas Kim, Giulio Zhou, Conglong Li, Hyeontaek Lim, David G. Andersen, Michael Kaminsky, Subramanya R. Dulloor. Conference on Machine Learning and Systems (MLSys), 2018.

## Teaching Experience

---

**Machine Learning Systems Seminar**

*Led paper discussions and held weekly office hours.*

CMU, Spring 2022

**Introduction to Machine Learning, PhD**

*Mentored projects, created assignments, and gave recitations.*

CMU, Spring 2020

**Introduction to Machine Learning**

*Created assignments and led weekly discussion sections.*

UC Berkeley, Fall 2016

**Data Structures and Algorithms**

*Created assignments and led 12 hours of lab instruction per week.*

UC Berkeley, Summer 2016

**Data Structures and Algorithms**

*Held office hours, created assignments, and led weekly discussion and lab sections.*

UC Berkeley, Spring 2016

## Honors & Awards

---

**NSF Graduate Research Fellowship Program, Honorable Mention.**

*Received an honorable mention in the field of machine learning.*

**EECS Honors Degree Program.**

*Honors program with 20-30 students. Requirements include research and studies in concentration outside EECS.*

**Tau Beta Pi, Engineering Honor Society.**

*Accepts the top 1/8th of students in the College of Engineering.*

## Technical Skills

---

**Programming Languages:** Python, Java, C, C++

**Software/Frameworks:** Tensorflow, PyTorch, OpenCV, Apache Beam