Jitian Zhao

Email : jzhao326@wisc.edu
Mobile : +1-608-886-8204

Education

- **University of Wisconsin-Madison** — Madison, WI
  Ph.D. in Statistics; GPA: 3.91 — *Sep. 2020 – present*
  Master of Science in Data Science; GPA: 4.00 — *Aug. 2019 – May. 2020*
  Visiting Student; GPA: 4.00 — *Aug. 2018 – May. 2019*
- **Zhejiang University** — Hangzhou, China
  Bachelor of Science in Statistics; GPA: 3.71 — *Aug. 2015 – Jun. 2019*

Work and Research Experience

- **Applied scientist intern** — May. 2023 - Sep. 2023
  Amazon Alexa AI — *Bellevue, WA*
  - **Multimodal vector search system**: Built a vector search system for multimodal data including text, image and video. Conducted comprehensive evaluation on embedding model, indexing method and indexing platform.
  - **LLM-augmented search system**: Proposed a novel two-stage LLM-augmented search method to enable search both by user-defined metrics and semantic relevance. Experimented with SOTA large language models.
- **Applied scientist intern** — May. 2022 - Sep. 2022
  Amazon Alexa AI — *Bellevue, WA*
  - **Cascade multi-task language model**: Proposed a novel transformer-based cascading multi-task learning framework for label auto-generation. Achieved 30% increase in relative accuracy for hard tasks compared to traditional MTL language models.
  - **Applications**: Improved labeling efficiency using Cascade MTL model by providing adaptive label hints based on previous task label selection. Built an interactive labeling API product and won top 10% in an organization-wide hackathon.
- **End-to-end structured prediction error analysis** — Sep. 2022 - present
  Advised by Prof. Fred Sala — *Madison, WI*
  - **Two-stage structured prediction algorithm**: Developed a two-stage structured prediction algorithm which incorporates an embedding algorithm to learn complicated label space geometries, followed by geodesic regression that leverages both the learned embedding and given features
  - **End-to-end error analysis**: Provided an innovative end-to-end error analysis for structured prediction task by establishing an error rate guarantee for the embedding algorithm and generalization error analysis.
- **Citation graph analysis on semantic scholar dataset** — Sep. 2020 - Dec. 2020
  Advised by Prof. Karl Rohe — *Madison, WI*
  - **Journal clustering**: Aggregated the citation network of 220 millions papers by journals. Applied VSP (vintage sparse PCA) on the citation graph and obtained journal field membership based on citation patterns.
  - **Statistical paper analysis**: Focused on the Statistics-related journal cluster and zoomed into paper-wise citation graph. Applied VSP and bff (best feature function) to obtain clustered statistical methods.
  - **Trend analysis**: Selected statistical methods of interest and tracked their development over time based on the appearance in papers' abstract.

Skills

- **Software**: Proficient in Python and R, experienced in SQL, Matlab and C
- **Packages**: Pytorch, Transformers, Networkx, Geomstats, Pandas, Numpy
- **Related Experience**: Co-authored book chapter *Transformers–Improving Natural Language Processing with Attention Mechanisms*, Sebastian Raschka, Hayden Liu, and Vahid Mirjalili. Machine Learning with PyTorch and Scikit-Learn, Birmingham, UK: Packt Publishing, 2022. ISBN: 978-1801819312.

Honor and Awards

- **2015-2016**: Zhejiang University Academic Excellence Award
- **2017**: Winner of the 15th Statistical Modeling Contest at Zhejiang University
- **2018-2019**: Exchange & Visiting International Student Academic Excellence Award at University of Wisconsin-Madison (twice)