Jason (Junjie) Zhu

jasonjunjiezhu.com

Email: junjie.zhu.jason@gmail.com
Mobile: 650-285-7123

Summary

I am a curiosity-driven and scientifically-trained builder with experience in AI/ML, Statistics, and Graph Algorithms. Drawn to hidden patterns, scalable impact, and high-agency teams, I have continuously been applying my skills to real-world problems: multi-modal RAGs, search products, biomedical discovery, etc.

Education

- **Stanford University** — Stanford, CA
  Ph.D. in Electrical Engineering · M.S. in Statistics — *2014 – 2020*
- **Franklin W. Olin College of Engineering** — Needham, MA
  B.S. in Electrical and Computer Engineering — *2010 – 2014*

Experience

- **Nexa AI** — Cupertino, CA
  Head of AI/ML — *Feb 2025 – Present*
  - **Leadership**: Leading a lean and fast-paced team to accelerate Gen-AI edge inference on any device.
  - **Local RAGs**: Developing privacy-preserving RAGs with small AI models and on-device vision capabilities.
  - **Agentic Systems**: Researching action-driven applications with new AI protocols (e.g., MCP, A2A).
- **Apple** — Cupertino, CA
  Machine Learning Engineer — *Jan 2020 – Feb 2025*
  - **Synthetic Data Generation**: Invented methods to test model robustness via high-dimensional perturbations.
  - **Preference Learning**: Designed cost-efficient offline A/B testing to handle user distribution shifts.
  - **System Evaluation**: Implemented production pipelines to evaluate query understanding and ranking systems.
- **Stanford University** — Stanford, CA
  Research Assistant — *Sep 2016 – Feb 2020*
  - **Graph Visualization**: Developed graph visualizations to interpret and analyze the Gene Ontology.
  - **Unsupervised Learning**: Created dimension-reduction methods for stem cell and cancer model systems.
- **Illumina** — San Francisco Bay Area
  Deep Learning Scientist (Internship) — *Jun 2017 – Aug 2017*
  - **Model Architecture**: Combined CNNs, RNNs, and ResNets to improve accuracy for base-calling applications.
- **10X Genomics** — Pleasanton, CA
  Data Scientist (Internship) — *Jun 2016 – Aug 2016*
  - **R/Python Pipelines**: Built and productionized pipelines for exploratory single-cell analysis.
- **Olin College of Engineering** — Needham, MA
  Research Assistant — *Sep 2010 – May 2014*
  - **Graph Theory**: Solved distance-2-based graph coloring problems for special graph families.
  - **Information Theory**: Modeled wireless networks with stochastic geometric and interference models.

Selected Publications

- Automatically Authoring Regression Tests for Machine-Learning-Based Systems. *ICSE, 2021*
- Progenitor identification and SARS-CoV-2 infection in human distal lung organoids. *Nature, 2020*
- Exploratory gene ontology analysis with interactive visualization. *Scientific Reports, 2019*
- Visualization and analysis of sc-RNA-seq data by kernel-based similarity learning. *Nature Methods, 2017*

See Google Scholar for full list::