# Jonathan Yin

jonathan.yin@yale.edu | jonathanyin.com | jonathanyin12 | jonathan-yin

## Education

### Yale University

*Sept. 2020 - May 2025*

Double major in Computer Science and Statistics & Data Science, GPA: 3.98/4.0

- Selected coursework: Distributed Systems, Parallel Programming, Randomized Algorithms, Deep Learning, Intermediate Machine Learning, NLP, Algorithms, Data Structures, Financial Economics, Linear Models, Probability & Statistics Theory, Discrete Math
- Took leave of absence after junior year to work on startup with funding from Y Combinator

## Experience

### Lifelike (YC S23)

*May. 2023 - Nov. 2024*

Co-founder — *San Francisco, CA*

- Took a leave of absence from Yale to build Lifelike with $500k funding from Y Combinator, a top startup accelerator
- Launched one of the first real-time low-latency AI phone calls, and scaled to over 300,000 total users with Next.js, FastAPI, AWS ECS, Elasticache, and MySQL
- Served high-volume production traffic from self-hosted LLMs with vLLM and SGLang and built a custom image generation inference library, incorporating various techniques like regional prompting, IP Adapter, and ControlNet
- Fine-tuned LLMs and diffusion models (like Llama 3 and FLUX) to build the first real-time visual roleplay platform with character/scene consistency

### Benchling

*Jun. 2022 - Aug. 2022*

Software Engineering Intern — *San Francisco, CA*

- Integrated chemical editor into the electronic lab notebook, allowing users to design molecules or chemical reactions within notebook entries
- Created endpoints to convert finalized chemical structures from notebook entries into registered entities usable across the platform
- Feature released to enterprise customers as part of September 2022 release

### Octant

*Jun. 2021 - Aug. 2021*

Machine Learning Research Intern — *Emeryville, CA*

- Used graph convolutional networks for molecular property prediction to determine efficacy of drug synthesis pipeline
- Built similarity search tool to optimize which products to synthesize for secondary screening rounds based on hits from primary screen
- Applied K-means and UMAP to developed tool to visualize, cluster, and interactively explore high-dimensional molecular features

### Broad Institute of MIT and Harvard - Regev Lab

*Jan. 2019 - Dec. 2020*

Machine Learning Researcher — *Cambridge, MA*

- Worked on improving GPCR binding prediction with compressed sensing, Bayesian methods, and machine learning
- Developed novel deep learning architecture to create more meaningful latent molecular representations
- Paper accepted and selected for oral presentation at 2020 NeurIPS workshop, Learning Meaningful Representations of Life

## Conferences

### Learning Meaningful Representations for Life

*Dec. 2020*

Neural Information Processing Systems (NeurIPS) 2020 workshop

- **Yin J***, Chung H*, Regev A. *A multi-view generative model for molecular representation improves prediction tasks* (paper)
- Combined multimodal representation learning with variational autoencoders to improve latent molecular representations (talk)

## Skills

**Languages** Python, TypeScript, C/C++, Go, R, Java, HTML, CSS

**Libraries/Frameworks** Next.js, FastAPI, Flask, TensorFlow, PyTorch, NumPy, Pandas, Matplotlib