

JOSH SANYAL

☎ 408-816-5343 | ✉ jsanyal@stanford.edu | 🌐 joshsanyal | 🌐 josh-sanyal

EDUCATION

Stanford University

2021–2026

B.S. in Computer Science (Systems), Intended M.S. in Computer Science (AI)

GPA: 4.09/4.0

- *Relevant Coursework:* AI (Machine Learning, NLP with Deep Learning, Systems for ML), Systems (Operating Systems, Parallel Computing), Theory (Modern Algorithmic Toolbox, Theory of Computation, Discrete Mathematics)
- *Orgs:* CS Section Leader, TreeHacks Organizer, Neo Scholar Finalist, Pear Garage, South Asian Society, Bhangra

EXPERIENCE

Co-Founder

Ambora Labs (YC W24)

01/2024–07/2024

- Built product analytics for LLM applications (evaluated LLM outputs, segmented user inputs, logging, dashboards)

Quantitative Trading Intern

Jane Street

05/2023–08/2023

- Analyzed market data, trained statistical models, developed trading strategies, built trading intuition

Software Engineering Intern

Meta

06/2022–09/2022

- Built a stream processing framework for realtime detection and filtration of noisy data (bots, click farms) and redundant data to reduce computational burden on downstream ML models
- Performed data analysis to evaluate time and performance benefits of different filtration schemes

Research Intern

Mayo Clinic Banerjee Lab

06/2019–10/2021

- Trained a sequential deep learning model to predict breast cancer recurrence using EHR data, proposing novel weak supervision methods to overcome severe class imbalance and improve performance by 12%
- Developed a model to detect insulin pump failure in clinical notes to automate surveillance of adverse medical events

Research Intern

Stanford Quantitative Imaging and AI Lab

06/2018–08/2019

- Developed an automated pipeline that takes in multi-parametric MR images of prostate tumors, aligns the images in the same image-space using shape-based registration, and detects prostate tumors and quantifies aggressiveness

SELECTED PUBLICATIONS

- **Sanyal, J.**, Rubin, D., & Banerjee, I. (2022). A weakly supervised model for the automated detection of adverse events using clinical notes. *Journal of Biomedical Informatics*, 103969.
- **Sanyal, J.**, Tariq, A., Kurian, A. W., Rubin, D., & Banerjee, I. (2021). Weakly supervised temporal model for prediction of breast cancer distant recurrence. *Scientific Reports*, 11(1), 1-11.
- **Sanyal, J.**, Banerjee, I., Hahn, L., & Rubin, D. (2020). An Automated Two-step Pipeline for Aggressive Prostate Lesion Detection from Multi-parametric MR Sequence. *AMIA Joint Summits on Translational Science*, 552–560.

PROJECTS

Digital Twin for Cell Therapy Manufacturing

- Worked with a series B cell therapy startup to provide predictability to the manufacturing process

Optimized nanoGPT

- Implemented and analyzed methods including quantization, pruning, KV caching, speculative decoding

TECHNICAL SKILLS

Languages: C/C++, Python, Java, HTML/CSS, JavaScript, MATLAB, SQL

Frameworks/Libraries: PyTorch, CUDA, React, Gensim, NLTK, Matplotlib

ACHIEVEMENTS

- International Science and Engineering Fair Finalist 2020
- 3rd place in Mathematics and Computer Science at National JSHS 2021
- USA Computing Olympiad Gold Division 2019
- 5-time AIME Qualifier (highest score 9) 2017–21