JOSH SANYAL

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

EDUCATION

Stanford University 06/2024 (expected)
B.S. in Computer Science (Artificial Intelligence Track), Minor in Mathematics GPA: 4.19/4.0

- *Relevant Coursework*: Programming Abstractions, Computer Organization and Systems, Discrete Mathematics, Machine Learning (Coursera), Natural Language Processing with Deep Learning (Audit)
- *Organizations*: TreeHacks Hackathon Organizer (Technology Team), Association for Computing Machinery

EXPERIENCE

Research Intern **Mayo Clinic's Banerjee Lab** 06/2019–Present

- **2019-20:** Developed encoded representations of unstructured clinical notes through weighted word embeddings
- **2019-20:** Trained a sequential deep learning model to predict cancer recurrence using EHR data, proposing a novel weak supervision methods to overcome severe class imbalance and improve performance by 12%
- **2020-21:** Automated the post-market surveillance of adverse events using clinical notes by training and externally validating a machine learning model to detect cases of insulin pump failure as a test case
- **2020-21:** Developed a model to detect mass effect from CT head reports for mortality prediction

Research Intern **Stanford Quantitative Imaging and AI Lab** 06/2018–08/2019

- Developed an automated pipeline that takes input multi-parametric MR images, segments the prostate gland, and aligns images in the same image-space using shape-based registration
- Trained a pixel-level deep learning model that detects prostate tumors and quantifies aggressiveness with improved accuracy and explainability over previous state-of-the-art model

SELECTED PUBLICATIONS

- **Sanyal, J.**, Rubin, D., & Banerjee, I. (2021). 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

HomesteadHacks Website

- Created and designed the website with live schedule, embedded registration forms, workshop content, prizes
- Supported user accounts for solving algorithmic and CTF-style problems with a live leaderboard

Online Mafia Game

- Implemented the popular social deduction game, Mafia, with custom roles and win conditions
- Incorporated networking to host centralized multi-device games with chat-based communication

TECHNICAL SKILLS

Languages: C/C++, Java, Python, MATLAB, HTML/CSS, JavaScript, PHP, Octave
Frameworks/Libraries: TensorFlow, Keras, PyTorch, Gensim, NLTK, Matplotlib

ACHIEVEMENTS

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