JOSH SANYAL

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

EDUCATION

Computer Science, B.S. **Stanford University** 06/2025 (expected)
Relevant Coursework: Machine Learning (Coursera), Multivariable Calc, Linear Algebra, Discrete Math

EXPERIENCE

Bioinformatics Research Intern **Stanford Rubin Lab** 06/2018–Present

- Led 4 bioinformatics research projects (with 3 first-author publications):
  - **2018-19:** Automating pixel-level assessment of prostate cancer in multi-parametric MRI
  - **2019-20:** Predicting breast cancer recurrence, 1 year in advance, using EHR data
  - **2020-21:** Automating post-marketing surveillance of adverse events via clinical notes
  - **2020-21:** Detecting mass effect from CT head reports for mortality prediction
- Created computational models using machine learning, computer vision, natural language processing
- Collaborated with grad students, medical residents, faculty and presented at weekly lab meetings

RESEARCH/PROJECTS

Breast Cancer Recurrence Prediction

- Trained a sequential deep learning model to predict recurrences using longitudinal EHR data
- Developed encoded representations of unstructured clinical notes through weighted word embeddings
- Incorporated weak supervision of unlabeled data using NLP-curated non-perfect training labels to increase training data, overcome severe class imbalance, and improve performance by 12%
- Published first-author journal paper in Nature's Scientific Reports, May 2021
- Utilized Python, (*libraries*): Keras, TensorFlow, NLTK, Gensim, Matplotlib

Prostate Tumor Assessment

- 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
- Published first-author conference paper with American Medical Informatics Association, May 2020
- Utilized MATLAB, (*toolboxes*): Image Processing, Computer Vision, Deep Learning

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
- Utilized Java, Socket Programming (TCP/IP), Swing Graphics, HTML/CSS

TECHNICAL SKILLS

Languages: (*proficient*) C++, Java, Python (*prior exp)* MATLAB, HTML/CSS, JavaScript, PHP, Octave

ACHIEVEMENTS

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