

Joshua Lazaro

📍 Stanford, CA ✉️ jelazaro@stanford.edu ☎️ 832 561 1956 in [Lazaro-Bioinformatics](#)



Education

PhD	Stanford University , Biomedical Data Science	Sept 2023 - Expected: 2029 GPA: 3.86/4.0
BS	The University of Texas at San Antonio , Statistics and Data Science	Sept 2019 – May 2023 GPA: 3.9 / 4.0

Research and Work Experience

PhD Researcher , Stanford: Mordecai Lab and CA Department of Public Health	Stanford, CA June 2024 - Present
<ul style="list-style-type: none">• Built an automated probabilistic forecasting pipeline in Python using XGBoost Distribution to predict rare-event counts for the CA Department of Public Health.• Outperformed 21 models in the 2024 CDPH forecasting challenge by engineering a novel negative binomial regression model for distributional case count forecasting.• Enabled data-driven scenario planning for the CA Dept. of Public Health by delivering high-resolution probabilistic visualizations of disease forecasts.	
Data Science Intern , Medicaid and Medicare Office of Enterprise and Data Analytics	Washington, DC June 2023 - Sept 2023
<ul style="list-style-type: none">• Developed a county-level Generalized Linear Model to predict Medicare Advantage adoption, identifying key geographic and financial drivers across the U.S.• Validated model performance against held-out data, achieving high accuracy (MAE=0.01) to ensure reliable forecasting for federal policy frameworks.• Delivered insights on geographic and financial drivers of MA adoption to inform future forecasting frameworks for federal agents	
Amgen Scholar , Yale School of Medicine: Townsend Lab	New Haven, CT June 2022 - Sept 2022
<ul style="list-style-type: none">• Quantified the effect size of colorectal cancer gene mutations in 1015 patient exomes through <code>cancereffectsizeR</code>• Calculated gene level mutation rates through likelihood estimates for the expected number of mutations in a gene to attain the average selection advantage conferred upon mutations• Discovered the role of 46 gene variants in cancer progression by quantifying their statistical effect size using computational and evolutionary techniques.	

Projects

Leveraging Large Language Models for Adaptive Mental Health Training	Project Report 
<ul style="list-style-type: none">• Built and evaluated an AI powered training platform to upskill mental health counselors• Engineered failure mode injections, developed a flawed AI training detection curriculum, and developed quantitative metrics to evaluate counselor reliance on flawed AI outputs• Collaborated with an India based crisis hotline to benchmark synthetic patient agents and integrate our AI Copilot into a live clinical workflow	
AI Climate Forecasting Using Probabilistic Modeling	Project Report 
<ul style="list-style-type: none">• Led development of a data science pipeline for probabilistic temperature forecasting using distributional machine learning (XGBD, AutoBNN, GP), generating 10,000+ Monte Carlo simulations per timepoint and analyzed via quantile scores	

Honors and Awards

NSF Graduate Research Fellowship Program	Apr 2025
1st Place, West Nile Virus Forecasting Challenge, CA Dept of Public Health	Jan 2025
MAC3 PhD Fellow in Precision Health (In Honor of Paul Berg)	Jan 2024
2nd Place, Advanced Team – National Security Agency (NSA) Rowdy DataThon	Oct 2022