JAMES FU

925-875-8886 | jamesfup@gmail.com | linkedin.com/in/jamespfu | jamespfu.com

EDUCATION

University of Texas at Austin — Expected Dec 2025
M.S. in Data Science, GPA 3.9/4.0
- Relevant Coursework: Data Science Statistics, Natural Language Processing, Probability and Simulation

University of California, Los Angeles (UCLA) — Sep 2020 – Jun 2024
B.S. in Computational Biology, Data Science, GPA 3.74/4.0
- Relevant Coursework: Data Structures and Algorithms, Machine Learning, Probability and Discrete Math

SKILLS

Programming Languages: Python, SQL, R, C++, JavaScript, HTML, CSS
Frameworks: PyTorch, TensorFlow, React
Tooling and Systems: Git, Jupyter Notebook (Anaconda), Pycharm, Tableau, APIs, ETL, Linux
Libraries: pandas, NumPy, scikit-learn, Matplotlib, Plotly, R Tidyverse

EXPERIENCE

Zipursky Lab — Sep 2023 – Jun 2024
UCLA Undergraduate Research Thesis — *Los Angeles, CA*
- Developed an ML pipeline using K-means to cluster 9M+ cells based on MERFISH spatial coordinates and gene expression. Identified three astrocyte subtypes and used SVM (87% accuracy) to validate them against known neuronal layer markers, assessing whether astrocytes exhibited similar layer specificity.
- Applied statistical modeling with Bonferroni correction to analyze 1,100+ genes, found 104 astrocyte markers.

UCLA Semel Institute — Nov 2022 – Feb 2024
Student IT Technician — *Los Angeles, CA*
- Spearheaded procurement and rental of refurbished PCs, maintained inventory, performed hardware upgrades/RMAs, and troubleshot hardware, reducing e-waste and saving $10K+ annually.
- Imaged, encrypted, and deployed 100+ HIPAA-compliant PCs. Optimized network infrastructure for 150+ offices by utilizing batch scripts and Excel to track inventory and identify non-compliant hardware.

Silicon Valley Education Foundation — Jun 2022 – Aug 2022
Computer Science Innovate Teaching Assistant — *San Leandro, CA*
- Facilitated student learning in introductory Python concepts taught via Adafruit's Circuit Playground Express.
- Designed and led 10+ Python lessons on recursion, control flow, data structures, and GitHub deployment.
- Debugged code and prepared supplementary coding tutorials and challenges to reinforce OOP concepts.

UCLA Health — Jun 2021 – Aug 2021
Web Developer — *Los Angeles, CA*
- Improved the Emergency Medicine website used by 140+ residents by enhancing site structure and building 4+ pages with HTML, CSS, and JavaScript. Managed Drupal modules and updated SQL databases to maintain content and functionality.

PROJECTS

Equitable Healthcare Cost Modeling through Deep Learning — Apr 2024 – Present
- Built Faircare, a Python-Flask ML-powered healthcare cost modeling tool using React, surpassing 300+ visits in one month. Integrated synthetic data generation using MEPS dataset, increasing data volume by over 253%.

Enhancing Robustness in Natural Language Inference (NLI) Models | paper — Oct 2024 – Dec 2024
- Extended Swayamdipta et al. dataset cartography codebase to support ELECTRA-small, addressing the BERT-base model limitation and optimizing the training set to prioritize hard-to-learn examples.
- Improved NLI robustness by 8% F1 when evaluated on 5+ novel contrast sets created by modifying an SNLI subset with syntactic distractors generated using OpenAI's GPT-4o API.

LEADERSHIP

Biomedical Engineering Society at UCLA — Nov 2022 – Feb 2024
Build Team Project Manager — *Los Angeles, CA*
- Led a team of 35 undergraduates in a year-long project, developing and teaching weekly modules on Arduino (C++), Processing (Java), circuitry, and CAD to help teams develop a full-fledged pulse oximeter.