Jay Warrier

408-893-4984 | jay.warrier@duke.edu

Education

Duke University Durham, NC
BS in Computer Science, BS in Math, 3.9 GPA *Aug. 2024 – May 2028*

Notable Coursework

Data Structures and Algorithms, Linear Algebra, Mathematical Cryptography, Probability, Topology, Multivariable Calculus, Computer Architecture, Abstract Algebra (G), Machine Learning (G)

Technical Skills

Languages: Java, Python, JavaScript, HTML/CSS, R, SQL, TypeScript
Developer Tools: Git, VS Code, PyCharm, IntelliJ, Eclipse
Libraries: Pandas, NumPy, Matplotlib, Tensorflow, Pytorch, Scikit Learn, Igraph, OpenAI API, FastAPI, Pydantic, React

Experience

AI Document Analyzer | *Python, FAISS, OpenAI,* June 2025 – August 2025
- Built a full-stack AI document analyzer with FastAPI and FAISS, enabling semantic search and question answering over uploaded PDFs, DOCX, and TXT files.
- Implemented chunk-level embedding using Hugging Face models and stored vectors for retrieval-augmented generation (RAG) via OpenAI's GPT-4.
- Built a React.js frontend with TypeScript, using hooks for state management and TailwindCSS for responsive styling

Research Intern - Full Time - Duke University | *Python, Scikit-Learn, OpenCV,* June 2025 – August 2025
- Used OpenCV's connected components algorithm to identify histology images with multiple nerve profiles and classified the main trunk and branching structure, saving 8 hours of manual work per patient.
- Using color statistics and scikit-learn's Logistic Regression model, created a binary classification algorithm to identify errors in segmentation.

Research Intern - Part Time - Duke University | *Python, OpenCV, IGraph* September 2024 – June 2025
- Using 3D image data of the vagus nerve, a key target for electrical therapies, extracted fibers via OpenCV's connected components algorithm and connected them between frames according to area overlap metrics. Converted the network into an interpretable graph via igraph.
- Improved efficiency by 20x by identifying bottlenecks with cProfiler and snakeviz, and cleaned the graph by removing isolated nodes and patching segmentation errors.
- Given a JSON of fiber composition data at a specific frame, used the graph to backtrack and generate composition data for previous frames. This allows for more targeted stimulation, minimizing side effects.

Gene Expression Project | *R, Spectral Clustering* September 2022 – January 2024
- Implemented scikit-learn's Spectral Clustering algorithm to analyze gene expression. The algorithm achieved similar performance (across a variety of clustering metrics) to traditional algorithms (PCA, UMAP, TSNE) while being 2x as efficient. Published results to the Journal of High School Science

Fitmigo - AI Personal Trainer | *Python, Mediapipe* August 2023 – June 2024
- Retrieved body coordinates from video data via Mediapipe, and converted the coordinates to key joint angles.
- Used binary classification models to model form quality, and created a function to adjust target reps based on past performance.

Experience

Robotics Programming Lead January 2021 – March 2024
FRC Team 6962 *Mountain View, CA*
- Created and maintained the robot codebase of 10,000 lines of Java code via Github
- Educated other members of the team via peer mentorship and off-season workshops
- Using LimeLight, developed a vision system to automate specific tasks increasing their efficiency and success rate.

Activities/Societies

Duke Chorale (Tenor), Duke Opera Theater (Baritone), IM Basketball (3x3, 5x5)