

JACKSON NEWMAN

Redwood City, CA, 94065 | 650-649-8204 | jpnewman167@gmail.com | linkedin.com/in/jacknewman |
<https://github.com/JNewman-cell/> | jacksonnewman.netlify.app/

WORK EXPERIENCE

Company	Role	Location	Duration
AMD	Software Engineer Intern, Internal Development Tooling	San Jose, CA	June 2023 - September 2023
Shellie.us	Full Stack Software Engineer Intern, MERN Stack	San Francisco, CA	June 2022 - September 2022

AMD
Software Engineer Intern, Internal Development Tooling San Jose, CA
June 2023 - September 2023

- Accelerated Vivado constraint processing by 50% with a new C++ pattern matching function.
- Developed unit tests achieving 100% coverage to ensure performance and accuracy for Wildcard Matching.
- Reduced Vivado memory usage by 2% by refactoring code to utilize Tessil C++ Hash map package.
- Automated memory, encryption tests, and key upgrades using Python, cutting testing time by 50%.

Shellie.us
Full Stack Software Engineer Intern, MERN Stack San Francisco, CA
June 2022 - September 2022

- Developed a React-based hierarchical UI with dynamic modals and edit functionality for exhibits.
- Created REST API integrations for editing and saving exhibit data in the NoSQL backend database.
- Created reusable React components, improving maintainability and reducing development time by 30%.
- Used React and Redux for state management, enhancing data consistency and reducing errors by 20%.

PROJECTS

AI Model Demonstration Website <i>Flask, AWS, GCP, LLM, REST API</i>	
• Configured and deployed 3 AWS EC2 Linux VMs for AI model performance testing and benchmarking.	
• Benchmarked Llama.cpp and vLLM server with LLMPERF to determine the fastest LLM inference on CPU.	
• Improved Stable Diffusion performance with NNCF quantization, cutting image generation time by 80%.	
• Deployed an end-to-end Flask web app with Nginx to showcase OpenAI-like LLM inference via HTTP.	
• Developed a Flask API server capable of handling hundreds of image generation requests.	
TinyGPT <i>Python, PyTorch, Numpy, Self Attention, Transformers</i>	
• Implemented a cutting-edge language model using PyTorch, enabling the generation of text from input.	
• Utilized advanced NLP techniques inspired by GPT-2, including Transformers, for effective pattern capture.	
• Demonstrated expertise in training and optimization within the GPT-2 framework for text generation.	
Autoencoder Visual Classification <i>PyTorch, Numpy, Seaborn, Pandas, Matplotlib</i>	
• Trained autoencoder for precise letter recognition and transistor positioning on MNIST and MVTEC dataset.	
• Engineered efficient autoencoder architecture with optimized feature extraction for image classification.	
• Demonstrated 100% accurate image classification using threshold pixel-by-pixel differences.	
Wafer Map Failure Classification <i>PyTorch, Numpy, Seaborn</i>	
• Architected a 14-layer model classifying wafer types with > 90% accuracy using the WM-811k dataset.	
• Devised Augmentation layers for a noise-resistant network, slashing 80% of random chip failures.	
• Analyzed and visualized accuracy and loss throughout iterations to identify over- or under-training.	

EDUCATION

University of California Santa Barbara	Santa Barbara, CA
Bachelor of Science in Computer Engineering	September 2020 - June 2024
3.7 GPA, Dean's Honors, Relevant Coursework:	
Data Structures, Algorithms, Operating Systems, Applied Machine Learning and AI, Embedded Systems	

TECHNICAL SKILLS

Programming Languages: C++, Java, Python, C, SQL (Postgres), JavaScript, HTML/CSS
Frameworks: React, Node.js, Flask, Bootstrap, Material-UI, Storybook
Developer Tools: Git, Docker, Jira, Google Cloud Platform, Confluence, Perforce, NPM, Firebase, GitHub Actions
Libraries: pandas, NumPy, TensorFlow, PyTorch, Keras, Seaborn, matplotlib