Jacob Hong-En Wei

Irvine, CA jhwei@ucsc.edu 714-227-1611 jacob-wei jjacobhw

Education

University of California, Santa Cruz Expected Grad: June 2027
Bachelor of Science in Computer Science GPA: 3.85

- **Concentrations:** Machine Learning, AI Engineering, Software Engineering, Applied Data Science
- **Coursework:** Computing in Python, Object Oriented Programming, Data Structures and Algorithms, Computer Architecture, Probability Theory/Statistics, C Programming in Linux, Machine Learning, Artificial Intelligence, Software Engineering, Linear Algebra

Experience

Machine Learning Engineer *Palo Alto, CA*
Inference.ai *June 2025 – Present*

- Engineered and implemented RAG pipelines that reduced hallucinations by 20%, improved query response time by 16%, and eliminated manual intervention, providing more accurate and autonomous business intelligence insights.
- Applied LoRA-based fine-tuned on LLMs using Hugging Face Transformers, reducing loss and perplexity by 23% to improve response consistency across wider domains while optimizing parameter efficiency for more specialized tasks.
- Optimized vector database performance in ChromaDB, achieving a 10% improvement in storage and retrieval speed through query chunking, indexing, and metadata filtering while enabling dynamic updates without full retraining.
- Deployed MCPs, improving LLM modularity and reasoning capabilities through structured utility use, providing real-time predictions, and enhanced API integration.

Robotic Processing Automation (RPA) Intern *Orange County, CA*
Healthcare Practice IT *June 2024 – Sept 2024*

- Engineered invoice processing pipelines using UI Path to parse and validate around 100 business emails per week, extract/validate business data, and provide invoice and billing information while reducing manual entry errors by 60%.
- Worked with accountants to deploy an automated state tax calculator in AutoTask, streamlining tax payment procedures, eliminating the need for dedicated days of the month for tax filing.
- Enhanced UI Path processes with PowerShell scripts, improving autonomous capabilities and decreasing runtime by 15%, saving approximately 25-30 man-hours of administrative labor and 4 hours of maintenance time every month.

Projects

Personalized Student Recruiter Agent

- Developed a chatbot to answer questions about students' experiences, skills, and coursework.
- Enabled local LLMs to parse and filter relevant chunks of documents, obtaining a response accuracy of 87.91%.
- Tools used: Python, Ollama, ChromaDB, Langchain, Hugging Face Transformers, RAG

Email Address Validator

- Designed an automated email parsing solution in UI Path, integrating with NeverBounce, eliminating manual operation for improved efficiency, and increasing the accuracy of faulty email detection by 30%.
- Tools used: Python, VBA, Computer Vision, Regex, UI Path, RPA

Customer Database

- Developed a UNIX-style database in C, implementing custom hashing to minimize collisions for efficient storage, lookup, retrieval, and deletion.
- Implemented command-line interfaces to maintain and support a lightweight system distribution.
- Tools used: C, Hashing, Linux, VIM

Skills

Programming: Python, C, C++, JavaScript, Typescript, HTML, CSS, RISC-V

Frameworks/Libraries: PyTorch, Next.js, React, Tailwind CSS, Pandas, NumPy, Scikit-Learn

Utilities: RAG, OCR, MCP, Regular Expressions, Vector Databases

Technical: Data structures/Algorithms, Machine Learning, Linear Algebra, RPA, Software Engineering, Microsoft Excel

Platforms: Windows, Linux/Ubuntu, VIM, Bash, Visual Studio, PowerShell

Interests: Cooking, Snowboarding, Golf, Tennis, Lifting, PC/Keyboard Assembly, Music

Languages: English, Chinese