Amanda Chan

Email: amandachan.tech@gmail.com ·Github·Linkedin

EDUCATION

University of Pennsylvania Aug 2023 - May 2025
M.S. in Engineering in Data Science (Department of Computer and Information Science)
GPA: 3.83/4.00
Courses: Machine Learning, Natural Language Processing, Deep Learning, Data Structures & Algorithms, Database Management

Hong Kong University of Science and Technology Sep 2019 - Jul 2023
B.S. in Physics
Honors: HKSAR Government Scholarship Fund, Lee Hysan Overseas Scholarship, Dean's List

SKILLS

Programming Languages: Python (LangChain, LlamaIndex, PyTorch, TensorFlow, Pandas, PySpark), JavaScript (React), SQL
AI Agent & ML Skills: AI Agents (LangChain, RAG pipeline, Multi-tool orchestration), Large Language Models (ChatGPT, LLaMA, Gemini), Deep Learning (Transformers, RNN, LSTM)
Platforms & Tools: GitHub, Docker, Google Cloud, Amazon Web Services, Streamlit, OpenAI API, MySQL, NoSQL

EXPERIENCE

AI Engineer Jun 2025 – Present
GitHub Repository: github.com/amandachantech/ai-agent-multitool United States
[AI Agent | LangChain | RAG | OpenAI API | Streamlit | Python Modular Design]
AI Multi-Tool Agent for Document, PDF, and CSV QA

- Designed and developed a **multi-tool AI agent** supporting natural language queries across text, PDF, and CSV formats, powered by LangChain, OpenAI API, and a Retrieval-Augmented Generation pipeline.
- Implemented **modular agent architecture** for scalable reasoning, context-aware retrieval, and automatic tool routing based on query intent.
- Built an **interactive Streamlit web app** with integrated visualization (bar, line, scatter charts) to present structured data analysis results.
- Engineered **robust API integration** with error handling and response optimization to improve reliability and user experience.

Machine Learning Engineer Intern Aug 2024 – May 2025
Wharton Research Data Services United States
[LLM | RAG | Langchain | OpenAI API | Chatbot]
Optimizing Retrieval-Augmented Generation (RAG) for Financial Document Analysis

- Developed a **Retrieval-Augmented Generation (RAG)** system using **Langchain**, **OpenAI**, and **Gemini APIs** to answer finance-related queries from **SEC filings**, including 10-K and 10-Q reports.
- Designed and executed experiments comparing embedding models (BCEmbedding vs. OpenAI Embeddings) and retrieval methods (BCEReranking vs. Reciprocal Rank Fusion) to optimize system performance.
- Enhanced **retrieval accuracy** from 21% to 50% and **answer accuracy** from 16% to 42%.
- Researched and implemented **table-based QA** techniques—including **TableRAG** and **text2SQL** approaches—boosting table-based answer accuracy from **32% to 97%.**

Image-Based Recipe Generation and Optimization with LLM (Capstone Project) Nov 2024 – Dec 2024
University of Pennsylvania United State
[Multimodal | Vision-Language Model | Computer Vision | Natural Language Processing | Model Fine-Tuning]

- **Developed an automated recipe generation system** integrating computer vision (CV) and natural language processing (NLP) to generate structured recipes from food images.
- **Enhanced baseline (InverseCooking) with prompt engineering and fine-tuning**: Used microsoft/git-base-coco for food descriptions, refined recipe fluency with TinyLlama-1.1B-Chat (Extension 1, score: 0.398), and fine-tuned it on 400 image-recipe pairs for more structured outputs (Extension 2, score: 0.418).
- **Implemented an LLM-based evaluation framework** assessing content similarity, structure, clarity, and usability.
- **Fine-Tuned LLM outperformed the strong baseline** (0.326) by **28.2%**, demonstrating significant improvement in recipe quality.