

RYAN TSAI

Machine Learning Engineer

San Mateo, CA • (925) 819-0887 • tsai.ryanw@gmail.com • [GitHub.io](#) • [LinkedIn](#)

MACHINE LEARNING PROJECTS

FastAPI RAG Book Recommendation Chatbot

Dec 2024

- Developed and deployed a book recommendation chatbot using a retrieval-augmented generation (RAG) pipeline (Qdrant vector database, OpenAI embeddings/completions). Built the backend (FastAPI) and frontend (Streamlit) in separate Docker containers and deployed to AWS EC2 via Docker Compose. Backend retrieves book descriptions and metadata based on user queries and generates conversational responses. Maintained per-user conversation history in memory for session-based interactions.

Clinical Dialogue Summarization

Oct 2024

- Fine-tuned an encoder-decoder transformer with LoRA for clinical dialogue summarization (in Hugging Face), achieving 97%/47% ROUGE-L on train/validation sets. Embedded samples with SentenceTransformer and visualized clusters to uncover dataset limitations (small size, noisy labels) that prevent generalization.

Histopathology Image Classification

Aug 2024

- Trained CNNs and vision transformers in PyTorch for histopathology image classification, achieving 94% AUC and 84% F1 score on the test set – outperforming the published benchmark. Reduced overfitting by applying transfer learning, geometric image augmentation, and custom resampling techniques adapted from tabular methods.

PROFESSIONAL EXPERIENCE

BEAM DATA

Toronto, ON

Data Science Consultant

Feb 2025-Present

- [\[Repo\]](#) Developed a vector indexing and content tagging pipeline (in [Haystack](#)) for university teaching materials, with custom components for parsing, metadata normalization, LLM-based tagging (OpenAI), and deduplication. Designed the Zilliz vector database schema, chunking algorithm, and tagging algorithm, optimizing the few-shot prompts and filtering noisy tags. Created labeled datasets to evaluate performance (93% recall, 78% precision). Created shell scripts to streamline Docker-based AWS Lambda deployment with embedding model caching and customizable options via environment variables.

QUALCOMM

Santa Clara, CA

Staff Modem Systems Engineer

May 2022-Dec 2023

- Led the development and deployment of a novel power amplifier calibration algorithm, reducing power consumption by 20% and calibration time by 10%. Analyzed and visualized calibration data in Python to inform algorithm design across 70+ 4G/5G bands. Designed the validation plan to identify/fix bugs and quantify performance gains. Resolved all signal quality issues raised by deployment test teams.
- Led the development and deployment of a new software feature to adjust the timing of transmit data captures (used for signal quality monitoring) based on characterized delay, eliminating all timing failures raised by deployment test teams. Designed the delay characterization plan and confirmed timing improvements by collecting validation data.

ZEKU

Palo Alto, CA

Senior Staff Modem Systems Engineer

May 2020-May 2022

- Designed a hardware-efficient digital notch filter in MATLAB that improved received signal quality by filtering single-frequency noise. Developed a noise vector estimation algorithm that reduced average convergence time by 300%. Tuned filter hyperparameters to optimize simulated performance across 10+ channel configurations and a 50 dB range of noise levels.

- Designed hardware-efficient upsampling/downsampling filter chains in MATLAB for data transmission and signal quality monitoring. Tuned system architecture and filter hyperparameters to minimize hardware complexity while meeting signal quality and noise rejection targets in simulation across 20+ sampling rate configurations in simulation. Applied linear regression to optimize filter parameters.
- Led end-to-end data transmission simulations in MATLAB, integrating digital and RF models. Designed validation plan and implemented standards-compliant performance metrics. Tuned model hyperparameters to meet signal quality and noise rejection targets across 10+ channel configurations.

QUALCOMM

San Diego, CA

Senior RF Systems Engineer

Nov 2017-May 2020

- Led the deployment of a novel TxIIP2 calibration algorithm that improved received signal quality by reducing second-order transmission noise by 40 dB across 30+ 4G bands. Analyzed calibration data in Python and MATLAB to develop a calibration sharing strategy that reduced calibration time by 50%. Resolved all issues raised by deployment test teams and customer engineering.
- Developed a Python automation tool to analyze the impact of circuit nonlinearities on transmission signal quality across 70+ 4G/5G bands, replacing manual Excel workflows. Eliminated human error, increased test coverage by over 1000%, and improved accuracy by up to 20 dB through power spectral density simulation.
- Owned the delivery of 4G/5G transmitter specifications for two chipsets, performing link budget analysis and collaborating with hardware and systems teams to optimize cost-performance trade-offs. Resolved all transmitter-related deployment issues and RMAs, contributing to two successful product launches.

QUALCOMM

San Diego, CA

Senior RF Hardware Engineer

Apr 2013-Nov 2017

- Owned the collection and delivery of 2G/3G/4G test data on internal deployment platforms. Optimized hardware configurations, settings, and calibrations to meet target performance metrics and resolved all hardware-related issues, contributing to multiple successful product launches.

EDUCATION

WECLOUDDATA

Toronto, ON

Machine Learning Engineering Program

2024

UNIVERSITY OF CALIFORNIA, LOS ANGELES

Los Angeles, CA

Master of Science, Electrical Engineering

2011-2013

UNIVERSITY OF CALIFORNIA, BERKELEY

Berkeley, CA

Bachelor of Science, Electrical Engineering and Computer Sciences

2007-2011