

Andrew Langdon

andrewlndn@gmail.com - 214-695-9608 - linkedin.com/in/andrew-langdon-583b1952

ML Engineer with 10 years in software, now focused on deploying ML systems, automating MLOps workflows, and building domain-specific LLM apps in production.

SKILLS

Programming Languages: Python, SQL

Machine Learning & Data Science Libraries: Large Language Models (LLMs), PyTorch, scikit-learn, Transformers, pandas, NumPy, Matplotlib, OpenAI API, LoRA

ML Engineering & Workflow Tools: Docker, Airflow, MLflow, FastAPI, JupyterLab, Jenkins

Data Infrastructure & Cloud Platforms: Google Cloud Platform, Pub/Sub, PostgreSQL, Kubernetes

ML & AI Concepts: Retrieval-Augmented Generation (RAG), Fine-tuning, Agent-based Systems

EXPERIENCE

Machine Learning Engineer at Kindred K12

Apr 2024 - Present

- Designed and developed an LLM-based autonomous web-browsing agent framework that allows developers to define agents that intelligently crawl and analyze web content. Implemented stochastic search strategies to explore and rank web links using ranking functions.
- Extended the browsing agent framework to browse, analyze, and index user-submitted URLs, including discovering and storing site-wide and item-specific copyright information.
- Engineered a semantic search system that enables teachers to efficiently discover primary sources by querying high-level topics (e.g., "Should freedom be sacrificed for national security?").
- Built an LLM-based system that evaluates topic gaps in the primary source catalog and deploys focused LLM agents to crawl the web for additional sources.

Co-Author of Zebra-Llama (Rare Disease LLM)

Jun 2024 - Dec 2024

- Co-authored *Zebra-Llama: A Context-Aware Large Language Model for Democratizing Rare Disease Knowledge*, focusing on Ehlers-Danlos Syndrome (EDS). The paper was published on arXiv: arxiv.org/abs/2411.02657.
- Created a fine-tuning dataset and RAG pipeline using 4000+ academic papers, improving citation accuracy from 52.3% to 70.6%.
- Fine-tuned with LoRA, improving thoroughness (+7.4%), accuracy (+4.2%), and clarity (+2.7%).
- Released the model as an open-source resource: huggingface.co/zebraLLAMA/zebra-Llama-v0.2

Senior Software Engineer at SupportLogic

Aug 2018 - Oct 2023

- Designed and built alerting system delivering millions of real-time ML signals to customers via Slack, email, and custom endpoints, leading to a 20% CSAT increase at Databricks.
- Deployed random forest model predicting customer escalations, reducing Salesforce's rate from 4% to 2%.
- Led refactor of alerting framework to horizontally scalable microservice architecture.

EDUCATION

Advanced Machine Learning Course

Dec 2023 - Jun 2024

- Intensive 20+ hours per week of study in deep learning, MLOps, GenAI, and applied ML.

Columbia University - New York, NY

2012 - 2014

- Computer Engineering, B.S., *magna cum laude*

Claremont McKenna College - Claremont, CA

2009 - 2012

- Economics and Engineering, B.A.

Project Lead, PatientFlowML (Capstone, Advanced Machine Learning Course)

Jan 2025 - May 2025

- Built a full MLOps pipeline for a patient readmission prediction model, exposed via a REST API for real-time inference.
- Deployed to Kubernetes with horizontally scaled inference and Airflow-orchestrated training and retraining via Jenkins.