

Luke Salamone 414-875-7293

github.com/lukesalamone | lukesalamone94@gmail.com | lukesalamone.com

Experienced Machine Learning Engineer proficient in search ranking, click prediction, large language models, and natural language processing. Open to full-time engineering roles.

PROFESSIONAL EXPERIENCE

Machine Learning Engineer, **Grammarly** August 2024 - Present

- Trained transformer-based T5 query denoising autoencoder, increasing retrieval coverage by 36%. Optimized inference code and deployed to Kubernetes cluster.
- Directed ML intern in end-to-end development of query intent classifier. Directed data instrumentation, architectural exploration, and production search inference optimization.
- Led cross-functional team to design and integrate two-tower writing action recommendation model. Assisted in offline evaluation and production architecture deployment.
- Developed pairwise knockout selection and iterative self-play techniques for synthetic data generation, enabling rapid experimentation.

Machine Learning Engineer, **TikTok** January 2022 - May 2024

- Increased click-through-rate by 2.1% and staytime 1.4% via introduction of efficient DCNV2 architecture in search ranking, while decreasing serving costs by 10%.
- Improved AUC on cold-start document CTR prediction by 0.8% over baseline via training of pointwise multitask learning model.
- Improved music retrieval HNSW index with click-query deep retrieval system, enabling a 15% nDCG improvement and 0.55% increase in overall music add rate in music search.

Machine Learning Engineer Intern, **Encyclopedia Britannica** Summer 2021

- Implemented state-of-the-art conversational chatbot utilizing vector-based knowledge retrieval. Chatbot now available to over 100 million monthly users.
- Improved website search results by 37% (MRR) through optimization of ML model.

Senior Software Engineer, **Capital One** February 2017 - February 2020

- Reduced ETL costs 90+% by building a high-volume Java library for data streaming.
- Reduced errors by 95+%, lead development of serverless app for data integrity / reporting.
- Designed data recommendation tool for internal metadata inquiries utilizing a hierarchical classification machine learning system.

TECHNICAL SKILLS

Programming:	Python, Java, Rust, C++, Javascript, SQL, Bash
Packages:	PyTorch, Tensorflow, Keras, Numpy, Pandas, Scikit-Learn, ONNX
Tools:	AWS, Azure, GCP, CI/CD, Version control (Git), Web Assembly (WASM)

EDUCATION

Master of Science (M.S.) - **Artificial Intelligence, Northwestern University** Fall 2021
Bachelor of Science (B.S.) - Computer Science, **University of Wisconsin-Madison**

PROJECTS

Chessmate AI – Trained Vision Transformer (ViT), achieving state-of-the-art chess move prediction (55% top 1). Fine-tuned for personalization via Direct Preference Optimization (DPO).

- Quantized and compiled to ONNX. Served at the edge to minimize inference cost.
- Pipeline written in Rust with Python bindings (PyO3). Additional Rust-WASM in browser.