

# Mike Lawrence

mike.a.lawrence95@gmail.com | LinkedIn: [www.linkedin.com/in/mlawrence95/](https://www.linkedin.com/in/mlawrence95/)

## SUMMARY

---

Machine learning engineer with experience in language modeling, recommender systems, and ML infrastructure. Enjoys hacking new things and big challenges. Always willing to help

## EDUCATION

---

UC Berkeley  
Physics, BA

December 2018

## WORK EXPERIENCE

---

### Google, YouTube

Senior Machine Learning Engineer

April 2024 – Current

Mountain View, California

**Summary: Techlead for flagship Shorts retrieval model**

- Planned and executed model quality iterations, driving millions of new engaged users and >1% watch time. Studies include model scaling, multimodal understanding, mixture-of-experts approaches, reward shaping, novel attention and position encoding schemes, data throughput optimization, and more
- Forged custom embedding initialization pipeline for TPU models, then used it to cheaply finetune pre-trained embeddings from large encoders to deliver high quality content to users
- Migrated large sequential recommender to latest TPUs for distributed training, optimizing input feed for 99% duty cycle. Helped other model teams do the same.
- Solved distributed serving bottlenecks, improving latency and serving duty cycle by >50% at scale
- Designed new offline evals, driving topline engagement wins via reduced training/serving skew
- Led collaborations between my team and verticals like Gaming, Creation, Shopping, and Ads to unlock further business value with my model

### Google, Cloud

Machine Learning Engineer

December 2021 – April 2024

Sunnyvale, California

**Summary: ML Training and Prediction pipelines on Vertex AI**

- Productionized Temporal Fusion Transformer into training and serving products, including attention-based feature importance and GPU-friendly serving container
- Found remote code execution vulnerability (RCE) on AI platform, then organized engineering response
- Rewrote dataloader for distributed training, optimizing for GPU throughput on large multimodal datasets

### Illumina

Data Scientist

May 2021 – November 2021

San Diego, California

**Summary: Machine learning engineer, financial forecasting platform**

- Improved financial forecasts with better probabilistic models and guards against temporal leakage

### Invoca

Senior Data Scientist

Jan 2019 – May 2021

Santa Barbara, California

**Summary: ML Engineer, real-time NLP on high-volume stream data**

- [Patented] Invented method for identifying robocalls with 90% accuracy using statistical language models
- Automated production deployment for a large C++ Bayesian graphical models codebase
- Created a self-serve tool for quickly interpreting live ML models

## RELEVANT SKILLS

---

**Languages:** Python, Java, C++, SQL, Bash

**Frameworks:** Pytorch, Tensorflow, Scikit-learn, Pandas, Numpy, Keras, Jax

**Infrastructure:** Docker, Kubernetes, Cloud computing, GPU/TPU computing and profiling