Robert Lewis

Email: jobs@robertlewis.dev

Los Angeles, CA robbylew robbylew Cell: +1 (323) 886-9730

Education

Georgia Institute of Technology (Georgia Tech) Remote

M.S. in Computer Science **Expected June 2026**

– **Relevant Coursework**: Database Systems Concepts and Design, Database System Implementation, System Design for Cloud Computing, Data Analytics and Security, Computer Networks

University of California, Los Angeles (UCLA) Los Angeles, CA

B.S. in Statistics and Data Science **Achieved June 2023**

– **Relevant Coursework**: Probability Theory, Deep Learning, Neural Networks, Bayesian Statistics, Computer Vision, Data Warehousing, Data Mining, Data Visualization, Data Structures & Algorithms

Technical Skills

Programming & Scripting: Python, Go, SQL, R, Bash

Machine Learning & Data Science: Pandas, NumPy, TensorFlow, PyTorch, scikit-learn, Matplotlib

Cloud & MLOps: AWS, Azure, GCP, Docker, Kubernetes, Terraform, Airflow, Jenkins, GitLab CI/CD

Data Engineering & Pipelines: Kafka, Hadoop, Spark, Hive, Jenkins, dbt, Apache Beam

Databases & Storage: MySQL, PostgreSQL, MongoDB, Redis, S3, Elasticsearch, Cassandra

Projects

Audify: Audio Fingerprinting & Song Recognition System GitHub Repo

– Engineered an end-to-end pipeline in **Golang** that re-implements Shazam's song recognition algorithm by extracting robust audio fingerprints.

– Integrated **Spotify** and **YouTube** APIs to fetch track metadata and download songs, while leveraging **FFmpeg** for audio preprocessing and conversion.

– Developed flexible data ingestion and matching processes using both **SQLite** and **MongoDB**, ensuring scalability and efficient retrieval of fingerprint data.

LLM-Driven Document Summarization Platform

– Aimed to streamline knowledge retrieval across **millions** of enterprise documents for faster decision-making.

– Engineered a scalable system by orchestrating **LLM** prompts with **LangChain**, leveraging **Pinecone** for semantic indexing, and containerizing the solution with **Docker**.

– Deployed on **AWS Fargate** for serverless scaling, reducing response latency by **30%**.

End-to-End MLOps Pipeline for Image Classification

– Implemented a full **Kubeflow**-driven pipeline handling data ingestion, model training, and validation; provisioned **GPU**-enabled **Kubernetes** clusters using **Terraform**.

– Reduced manual intervention by **40%**, cut inference latency by **30%**, and maintained cost-efficient scalability for large-volume image processing.

Certifications

CompTIA A+

CompTIA Network+