

# KELVIN CHRISTIAN

Software Engineer | Machine Learning Engineer

📞 206-747-9126    ✉️ christian.ke@northeastern.edu    🌐 kelvinc123    📧 kelvinc123    🌐 kelvinc123.com

## EXPERIENCE

### Statistical Analyst (Software/AI Engineering)

Nov 2024 – Present

County of San Bernardino

San Bernardino, CA

- Shipped **20+** internal tools including PDF parsers, browser automation, and scheduled ETLs, deployed across **4** departments
- Designed end-to-end pipelines for **7** programs with Playwright scraping, Pandas transforms, and automated ArcGIS refresh
- Created a dataset-discovery RAG app over the county data catalog to recommend sources and next steps for new initiatives
- Built ReAct agent (LangGraph and Google Search) with retries, token limits, and logging; made data enrichment **87%** faster
- Delivered a cost-optimized, fault-tolerant LLM pipeline for tabular data (row batching, parallel workers, checkpoint retries)

### Software Engineer - Generative AI

Feb 2024 – Nov 2024

Chekable

Oakland, CA

- Collaborated with frontend engineer and UX/UI designer to deliver backend features for a generative-AI patent platform
- Reduced end-to-end response time by **40%** by optimizing gRPC, Flask RESTful APIs, and MongoDB queries
- Built RAG pipeline with LLM and agentic patterns to generate **12+** page patent applications, reducing drafting time by **75%**
- Secured **\$150k** Azure startup credits and led infrastructure migration from Google Cloud to Microsoft Azure
- Developed a real-time collaborative editor and AI copilot using WebSockets, supporting **10+** concurrent users per project
- Designed Redis Pub/Sub architecture to synchronize state across **3** distributed server replicas for real-time applications

### NLP Research Intern

Jun 2023 – Aug 2023

Institute for Experiential AI

San Jose, CA

- Executed **50** parallel data processing jobs on HPC via Slurm and bash to efficiently handle **6 TB** of LaTeX and PDF documents
- Extracted and processed **1.7M+** table images and text from PDFs using Table Transformer and OCR models via Hugging Face

### Data Scientist Intern

Jul 2022 – Sep 2022

Cisco Systems

Seattle, WA

- Cleaned and analyzed **1.2 TB** of real-time DNS queries from AWS S3 using PySpark to improve DNS security analytics
- Cut ETL runtime by **70%** on Databricks by converting Parquet datasets to Delta tables and optimizing PySpark SQL queries
- Developed anomaly detection system using EM clustering to flag outliers among **100k+** users based on DNS request patterns

### Software Engineer Intern - Sole Engineer

Jan 2022 – Jun 2022

Aeyesafe

Seattle, WA

- Built **25+** REST APIs across **5** Express.js microservices with MongoDB database for real-time monitoring of **14** patient devices
- Trained/benchmarked pose-estimation models; best VGG-16 model reached **96%** test accuracy with a reproducible pipeline
- Implemented system health monitoring dashboard using InfluxDB and Grafana, streaming request status, latency, and errors

## PROJECTS

### iSEArch Hotels 🌐 | Python, Docker, Streamlit, Qdrant, LangChain, OpenAI

Mar 2024

- Won **2nd place** at Traversaal.ai Hackathon by building a RAG-powered travel recommendation chatbot using hotel datasets

### Diffusion-enhanced Mask Aware Transformer (DeMAT) 🌐 | PyTorch, TensorBoard, VAE, Diffusion, SAM

Dec 2023

- Integrated MAT with Stable Diffusion VAE for context-aware inpainting; improved reconstruction quality by **39.3%** SSIM

### Distributed E-Commerce Platform 🌐 | Node.js, Express.js, RabbitMQ, Docker, MongoDB

Mar 2023

- Built fault-tolerant microservices backend using async messaging and service isolation for horizontal scalability

## EDUCATION

### Stanford University

Artificial Intelligence Graduate Certificate

Online

In Progress

### Northeastern University

M.S. Computer Science, GPA: 3.88/4.00

San Jose, CA

2021 – 2023

## SKILLS

**Programming Languages:** Python, R, SQL, C, C++, Java, JavaScript

**Databases:** MySQL, MongoDB, Redis, PostgreSQL, Qdrant, Milvus, Weaviate, Elasticsearch, InfluxDB

**Machine Learning:** PyTorch, TensorBoard, Spark, OpenCV, NLTK, scikit-learn, statsmodels, Pandas, NumPy, Matplotlib, Pyomo

**Tools:** AWS, GCP, Azure, Databricks, Socket.IO, RabbitMQ, Docker, Airflow, OpenAI, LangChain, LangGraph, Cursor