Claire Lee

550 Battery St., San Francisco, CA 94111

415-361-0188 • glee2429@gmail.com • glee2429 • glee2429

Experience

Slickdeals **San Francisco, CA**

AI Engineer & Tech Lead *2023 – Present*

- Architected and deployed a high-performance data ingestion pipeline for ElasticSearch, integrating batch processing and a real-time event-driven Kafka system, enabling seamless search indexing for millions of records.
- Led the revamp of cloud deployment infrastructure, streamlining automation with Env0, Terraform, Helm, and AWS, reducing deployment time by 60%.
- Engineered a vector search solution by provisioning RDS instances and optimizing PostgreSQL with Pgvector and HNSW, improving nearest-neighbor search speeds by 5x for AI-powered recommendations.
- Designed, implemented, and scaled an AI-driven summarization service using GPT-4o, optimizing token usage to balance performance and cost-efficiency, reducing inference costs by 30%.
- Spearheaded AI adoption within the organization, bridging the gap between engineering and business stakeholders, leading AI strategy, and conducting technical workshops and demos to drive executive buy-in.

Kiva Microfunds **San Francisco, CA**

Software Engineer *2021 – 2023*

- Designed and implemented a high-performance deduplication system using DynamoDB, reducing redundant records and improving data consistency across financial transactions.
- Led a major data migration project, transforming and migrating terabytes of user data from a legacy system into a microservices-based architecture with zero downtime.
- Developed an asynchronous event-driven messaging system to decouple a monolithic checkout processor, improving system reliability and transaction processing speeds by 40%.
- Designed and implemented an optimized API contract to fetch data from ElasticSearch, reducing response times by 50% and enhancing search efficiency.
- Led continuous monitoring and performance tuning efforts, proactively identifying and resolving critical issues using DataDog and Confluent, minimizing API failure rates.

Data Management and Analytics Laboratory (OSU IDEA Lab) **Corvallis, OR**

AI Researcher & Database Developer *2019 – 2021*

- Led cutting-edge research on AI-driven database optimization, developing novel machine learning algorithms to improve query efficiency and data integrity.
- Designed and implemented scalable AI models for relational learning over large, noisy datasets, significantly improving data cleaning automation.
- Published a first-author research paper in SIGMOD 2020, one of the most prestigious peer-reviewed conferences in AI and Database Management Systems.
- Presented research findings to industry experts and academia, demonstrating innovative approaches to AI-powered data management.

Novartis Pharmaceuticals **Seoul, South Korea**

Data Scientist *2018 – 2019*

- Designed and deployed an AI-driven patient monitoring platform, leveraging predictive analytics to detect early signs of heart failure in accordance with the American Heart Association (AHA) guidelines.
- Integrated real-time patient health data with machine learning models, enabling automated alerts for at-risk patients and facilitating early medical intervention.
- Drove cross-functional collaboration between engineers, physicians, and regulatory teams to ensure compliance with medical standards while enhancing patient outcomes.

Kiva Microfunds **San Francisco, CA**

Data Analyst, Strategy Operations *2018*

- Developed and deployed a machine learning model to identify high-risk borrower profiles, reducing the fraud detection time by 40% and improving the efficiency of loan approval.
- Automated and optimized ETL pipelines to process large-scale transactional data, improving reporting speed, and reducing operational costs.
- Designed and maintained interactive data visualization dashboards, providing real-time insight to executives and streamlining decision-making.

Education

Master of Science in Computer Science and Artificial Intelligence **Corvallis, OR**
Oregon State University
GPA: 4.0 **Thesis:** Relational Learning over Dirty Data Using Data Constraints

Bachelor of Science in Applied Mathematics and Statistics **New York, NY**
Columbia University in the City of New York
Capstone Project: Machine Learning in Drug Discovery

Technical Skills

Programming Languages: Golang, Kotlin, Java, Python, SQL, JavaScript, C++, Shell Scripting, Bash
Machine Learning & AI: LangChain, OpenAI API, Vector Databases (Pgvector, FAISS), HNSW, LLM Optimization
Cloud & Infrastructure: AWS (S3, DynamoDB, RDS, EC2, Lambda), Terraform, Env0, K8s, Docker, Helm
Data Engineering: Dagster, Kafka, Airflow, ElasticSearch, Redis, PostgreSQL, DynamoDB, Snowflake
DevOps & Monitoring: CI/CD (Jenkins, GitHub Actions), Datadog, Confluent, Sumo Logic, Prometheus, Grafana