

Jerry

Data Engineer | Financial Data Infrastructure & Data Platforms
github.com/jungsooyun

Summary

Data engineer who has designed and operated data pipelines and platforms for large-scale recommendation systems at Kakao and for institutional financial data products at CryptoQuant. Experienced in dbt, Spark, Airflow, Databricks and Delta Lake, and in the Kubernetes infrastructure beneath them. Builds and runs a personal quant trading system on crypto derivatives and DeFi, and applies AI coding agents actively to engineering work.

Experience

CryptoQuant

Web3 Data Engineer

Aug 2023 – Present

Data platform modernization: open table format serving

- Migrated the PostgreSQL-centered API serving layer to Delta Lake on Databricks to meet the varied data consumption needs of institutional clients, including US brokerages
- Exposed an Iceberg catalog through Delta UniForm so clients can connect Snowflake and BigQuery directly, opening a data sales channel beyond the API

Analytics pipeline standardization and improvement

- Managed and optimized the data pipelines of the BigQuery-based analytics platform
- Standardized transformation pipelines with dbt and moved large-scale computation to distributed Spark processing, shortening data latency

XRP Ledger data product

- Built pipelines that collect XRP Ledger and market data into PostgreSQL and BigQuery
- Secured a competitive data edge with real-time updates and transaction data decoding

Workflow orchestration and infrastructure

- Built and operated HA Airflow managed with ArgoCD, Kustomize, and Helm, securing pipeline stability and operational efficiency
- Moved EC2 services to EKS and introduced Karpenter autoscaling; provisioned batch workloads quickly and flexibly with SPOT instances and Kubernetes restart logic

Kakao

ML/Data Engineer, Recommendation Team

Dec 2019 – Aug 2023

Integrated pipeline for high-volume user logs

- Developed and operated pipelines that integrate and preprocess high-volume user logs in real time across KakaoTalk, Gift, advertising, Daum News, and Daum Story
- Developed and maintained recommendation systems for multiple services, improving user engagement and experience

Led the adoption of Airflow 2.0

- Led the adoption of Apache Airflow 2.0, improving the scalability and operational efficiency of task orchestration (presented at if(kakao))

Search keyword ad recommendation pipeline

- Built a data pipeline that predicts pCTR from Daum Search query-consumption history and recommends keyword ads
- Reached responses within 50 ms with vector DB (ANN) serving

Education

Yonsei University

Bachelor's in Industrial Engineering

Feb 2013 – Aug 2019

Projects

Quant trading system: build and operation

Jan 2025 – Present

A personal project to design, implement, and run systematic trading strategies on crypto derivatives and DeFi

- Designed and ran a delta-hedged market-making strategy that combines Uniswap LP positions with futures positions
- Implemented synthetic strategies on futures and options, and automated position management and risk monitoring
- Built pipelines that collect and analyze on-chain and off-chain market data, used for signal generation and strategy backtesting

Activities

Decipher blockchain research group

Aug 2023 – Aug 2024

Study sessions, conference talks, and research articles

- Research on running Lightning Network nodes (YouTube, in Korean)
- Scaling solutions for monolithic chains (article, in Korean)
- Hands-on on-chain data analysis (in Korean and English)
- Interplanetary Money (article, in Korean)

Skills

Data Engineering & Analytics

Spark, Airflow, Python, SQL, dbt, Kafka, BigQuery, Databricks, Delta Lake / Iceberg

Infrastructure & Cloud

AWS, Kubernetes, ArgoCD, Helm, Karpenter, Prometheus & Grafana

Databases

PostgreSQL, MongoDB, Redis

Languages

Korean (Native) · English (Conversational)