Sung Kim

+1 (925) 858-5663 | San Ramon, CA | [LinkedIn](#) | [GitHub](#) | [Website](#)

EDUCATION & SKILLS

Stony Brook University   Stony Brook, NY
 B.S. in Applied Mathematics and Statistics   *Aug 2021 – May 2025*

Certificates: Databricks Certified Data Engineer Professional, AWS Certified Solutions Architect
Specialities: Data Pipelines, Cloud Computing, Unix/Linux, Scripting, Web development
Technologies & Tools: Python, JS/TS, Bash, SQL, Docker, Git, AWS, Airflow, Snowflake, Databricks

EXPERIENCE

Rocket Lawyer   San Francisco, CA
Data Engineer, Intern   *June 2024 – Oct 2024*
- Maintained production data pipelines using Airflow, Snowflake, and GCP, ensuring proper DataOps
- Engineered ELT workflows handling 5M+ records, implementing backfilling and data normalizations
- Architected and built efficient ELT pipelines using Airflow with Medallion architecture, enabling bronze-silver-gold data transformation layers
- Performed SQL analysis and optimization in Snowflake, reducing query runtime by 17% and improving pipeline efficiency
- Automated data lifecycle management using cron jobs and scripting, reducing cloud storage cost by 5%

BrainSoft.AI   Seoul, South Korea
Data Scientist, Intern   *May 2023 – Sep 2023*
- Conducted feature engineering of cryptocurrency data API using Python and Tensorflow
- Optimized a custom FFT prediction algorithm through error analysis, reducing prediction error by 3%
- Developed secure data management platform with CRUD operations, serving 10k+ customer records
- Deployed containerized REST API data pipeline using Docker and GCP serving data to the platform

HearstRnH   Montgomery, AL
Software Engineer   *Dec 2021 – Dec 2022*
- Effectively led a small team to develop scalable web application adhering to Agile SDLC
- Deployed mobile-first homepage as serverless architecture using Nginx, React, Express, and AWS
- Optimized AWS cloud infrastructure through performance monitoring, reducing operational cost by 13%

PROJECTS

Homelab   *July 2025 - CURRENT*
- Configured home network infrastructure with redundant Proxmox & Raspberry Pi 4 clusters
- Architected modern data collection and backup using Dagster, DBT, PostgreSQL, and S3 on AWS
- Secured inbound network connection from the internet with Cloudflare ZeroTrust tunnel

Movie recommendation API ([Github](#))   *Nov 2024*
- Performed distributed data processing and feature engineering with Hadoop on 1M+ MovieLens dataset
- Deployed movie recommendation API using Flask on AWS, serving with hybrid KNN/SVD model