

TE-KAI CHOU

Kelvin Chou | bestkelvin08@gmail.com | San Jose, CA | (470) 601-4801 | www.linkedin.com/in/tekaichou/ | github.com/TeKaiChou

PROFESSIONAL SUMMARY

Data/Analytics Engineer with 2+ years of experience in building scalable ELT/ETL pipelines and analytics solutions in cloud environments. Proven ability to collaborate cross-functionally with product, operations, and data teams to deliver data models, dashboards, and automation that drive business impact. Proficient in SQL, Python, dbt, and GCP, with strong interests in expanding into multi-cloud, workflow orchestration, and data observability tools.

WORK EXPERIENCE

Taelor | Fashion Tech Startup **Hayward, CA**
Jul 2023 – Present
Data Analytics Engineer

- Developed and orchestrated a suite of 10+ data pipelines using **Prefect** and **Cloud Run**, achieving **99.9%** data reliability and timelines for critical business reporting and analytics processes.
- Designed and implemented a scalable data warehouse solution **from 0 to 1** using **BigQuery** and **dbt**, building 20+ data models and reducing query times by **80%**, while accommodating a **5x** growth in data volume.
- Engineered a robust ELT framework using **Python** and **SQL** with incremental models, which improved data refresh efficiency by **60%** and reduced infrastructure storage costs by **25%**.
- Collaborated with cross-functional teams to design and implement 10+ interactive dashboards using **Looker Studio**, providing daily insights that drove a **20%** increase in operational efficiency across departments.
- Introduced automated data quality checks and monitoring systems, leveraging **Great Expectations** and **dbt test**, resulting in a **70%** reduction in data-related incidents and a **50%** decrease in mean time to resolution.

BlueX Trade | Logistics Fintech SaaS **Taipei, TW**
Jul 2022 – Aug 2022
Data Scientist Intern

- Forecasted merchant account liquidity using **SARIMA** models in **PyCaret** and **Plotly**, improving underwriting response time by **1 day** and enabling faster lending decisions.
- Built a core data warehouse leveraging **Snowflake**, **dbt**, and **SQL**, and orchestrated daily ETL pipelines with **Airflow**, cutting reporting latency by **50%** and increasing freshness KPIs.
- Collaborated with data engineers on SageMaker model deployment pipelines, using Docker and ECR to improve testing workflow and reduce model iteration time by **30%**.
- Optimized existing SQL queries and data models, resulting in a **30%** reduction in average query execution time and **25%** decrease in storage requirement for the data warehouse

PROJECTS

Weather Streaming Analytics

- Developed an end-to-end real-time streaming pipeline** using Dockerized **Kafka**, **Spark Structured Streaming**, and **ClickHouse**, enabling low-latency ingestion, windowed aggregations, and scalable analytics.
- Built an interactive real-time dashboard** in **Streamlit** connected to ClickHouse, providing city-level weather trends and near real-time visualization for continuous monitoring.

Conversational AI for PDFs

- Developed a **Q&A Conversational AI** for PDF information summarization and retrieval by leveraging **Retrieval Augmented Generation (RAG)** solution powered by **Google Gemini-Pro API**, **LlamaIndex**, **Streamlit**, and Custom Prompts.

EDUCATION

Georgia Institute of Technology Atlanta, GA
Dec 2022
Master of Science in Operations Research (Data Science Concentration)

National Chiao Tung University (NCTU) Hsinchu, Taiwan
Jun 2020
Bachelor of Science in Industrial Engineering & Management

SKILLS

- Programming Languages:** Python, SQL, R
- ETL/ELT Tools:** dbt (data build tool), Prefect, Airflow, Kafka, Spark
- Cloud/Data Platforms:** GCP (BigQuery, Cloud Run, GCS, etc.), AWS (S3, SageMaker, ECR)
- Data Visualization:** Looker Studio, Tableau, Power BI, Jupyter Notebook, Streamlit
- Data Analysis / Machine Learning:** Regression Analysis, Statistics, Time Series Analysis, KNN, K-means, Naive Bayes, A/B Testing, Natural Language Processing (NLP)
- Database/Data Warehouse:** MySQL, PostgreSQL, BigQuery, Snowflake, ClickHouse, Dimensional Modeling
- Tools and Technologies:** Docker, Git, Scikit-Learn

CERTIFICATIONS

- Data Scientist Associate** by DataCamp Nov 2024
- Career Essentials in Generative AI** by Microsoft and LinkedIn Nov 2024
- AWS Certified Cloud Practitioner** by Amazon Web Services Jan 2023