Willie Huang

tyc30827@gmail.com • Willie Huang • github.com/tyc30827

WORK EXPERIENCE

XRSpace Co., Ltd. | Taipei, Taiwan   Aug 2024 - Present

Senior Data Engineer / MLOps Engineer

- Designed and developed AI-labeling system using LLM (AWS Bedrock, PydanticAI), Flask, Kafka, and Dagster for automated content tagging
- Built AI agent with dual-layer memory system and optimized RAG knowledge base to **combating hallucinations** and **reducing system uncertainty**, achieving **100% improvement in context relevance (0.35→0.7)** and **39% improvement in coverage (0.7→0.97)**
- Engineered end-to-end Avatar data pipeline for Japan app launch in **2 sprints**, coordinating **7 cross-functional teams** with dbt-governed BigQuery transformations and Looker Studio dashboards
- Optimized BigQuery data warehouse (clustering/partitioning) and complex SQL queries, reducing query costs by **90%**
- Led data platform infrastructure (Airbyte, dbt, Dagster, Terraform) achieving full automation and mentored junior engineers as Code Owner

CyberLink Corp. | New Taipei, Taiwan   Oct 2023 - Aug 2024

Senior Data Engineer

- Built NLP model (LDA) for customer feedback analysis and visualization in Tableau
- Developed data pipelines with Airflow and Airbyte to collect mobile install data
- Migrated data infrastructure from MSSQL to AWS S3 and Athena

CyberLink Corp. | New Taipei, Taiwan   Jun 2020 - Oct 2023

Data Scientist and Engineer

- Built personalized recommendation system, improving CTR from **<1% to 30-60%**
- Developed GNN-based user preference system using AWS Neptune
- Implemented MLOps practices with MLflow for model lifecycle management

EDUCATION

National Chiao Tung University | Hsinchu, Taiwan   Sep 2017 - Aug 2019

Master of Science in Information Management

Publication: First Author in **Information Fusion** (Top-tier AI Journal, 50+ Citations) | Attentive gated graph sequence neural network-based time-series information fusion for financial trading

National Cheng Kung University | Tainan, Taiwan   Sep 2013 - Jun 2017

Bachelor of Science in Industrial and Information Management

SKILLS & CERTIFICATIONS

Programming Languages: Python, SQL, Java, C++/C

Data Engineering Tools: Airflow, Dagster, Airbyte, dbt (Data Build Tool), Apache Kafka, Terraform

Databases / Data Warehouse: PostgreSQL, MongoDB, BigQuery, AWS Athena, AWS S3, DuckDB, MySQL, SQL Server

Machine Learning & MLOps: PyTorch, Keras, MLflow, Databricks, AWS SageMaker, GNN, Recommendation system, Financial Time-Series, Huggingface

GenAI / RAG / MLOps: AWS Bedrock, AWS Bedrock Knowledge Bases, PydanticAI, LlamaIndex

Backend: Flask, Pydantic, FastAPI, Dramatiq

Infrastructure: Docker, Git, GitOps/ArgoCD

AWS Services: Bedrock, Personalize, Rekognition, Neptune, EC2, Lambda, Kinesis

Data Visualization & BI: GA4, Looker Studio, Superset, Tableau

Package dependencies management: uv, poetry

Certifications: TOEIC 845