



Connor (Yu-Ming Hsu)

A software engineer building real product and interested in writing.

  leafwind.cs@gmail.com

Summary

- 10 years experience of large scale AI product building, and is capable of building product from scratch.
- Extensive problem solving experience for data science/engineering, and familiar with transferring real problem into requirements and solution planning.
- A well-rounded engineer in data project who bridges the gap between scientists and engineers.
- A pragmatic and ownership driven person, experienced with gap analysis, migration plan and release management.
- Leader of documentation and process, mentor of junior engineers.
- Product-driven mindset, focusing the impact against the cost to align team priorities.

Skills

Data Lake/Warehousing: Python, Spark, Scala, Airflow, Presto | **Streaming:** Kafka, Flink

Machine Learning: Scikit learn, Tensorflow (side project), TensorFlow Serving

AWS Cloud Services: EC2/ECS/ECR/Lambda, EMR (Hive), RDS, DynamoDB, ElastiCache (Redis), Athena, SNS, SQS, Glue, SageMaker, CloudFormation/CDK.

API Services: FastAPI, Flask, Django, Swagger/Flasgger | **CI/CD:** Jenkins/CircleCI/TravisCI, Ansible

Human Languages: English: Business Level | Japanese: Conversational | Mandarin: Native

Experience

Data Engineer, SmartNews, 2024 - Present

Data Engineer, Data Platform, Sep. 2024

- Manage data stacks and infrastructures to serve the product in SmartNews.

Senior Software Engineer, Moneytree, 2020 - 2024

Senior Software Engineer I, Nov. 2023 - Jul. 2024

- Migrated source of Lakehouse without impacting external data products.
- Lead and build a new team, enable junior members to contribute within a month.

Software Engineer II, Nov. 2020 - Nov. 2023

- Built a Data Lakehouse solution on a full cloud environment (AWS).
- Involved with **3 new data products** which made initial revenue stream for a new team with only **2** engineers.

- Lead a cross team initiative to democratize company crucial KPI as our dog food for business insight.
- Enable unit test on Spark locally, integrated with CI pipeline.
- Migrated CDK to version 2, improved CI workflow.
- Evangelizing Big Data technics with talks and drive technical improvements discussion every week.

Software Engineer, SmartNews, 2019 - 2020

- Data platform team: build data infrastructure/platform to serve Ads/Data Science needs.
- Improve Hive performance to **4x** by utilizing partition **after 1 month**.
- **Tech lead**: Lead team documentation culture, manage the process, and attend as representative in cross-team meeting **after 2 months**, Mentor 3 new members to equip the new data team with the full speed.
- Streaming locational service data for local coupon service for millions of users.

Senior Engineer II, Appier, 2014 - 2019

Senior Engineer II, 2018 - Jul. 2019

- Data governance chair / Agile coach: Coordinating cross-team product features, burning down company level technical debts and mitigated them gradually with shipping new features.
- AI backend: Building ML / OLAP services with F2E, Data Scientists, and Data Infrastructure team.
- Mentor new members by brining Advertisement domain/technical knowledge and formalizing documentation.

Software Engineer, Jun. 2014 - 2018

- Build RTB bidding algorithm in a fast-growing, dynamic business environment.
- Conduct experiments on real product to make daily improvements and achieve business goals.
- Conduct root cause analysis and solve critical issues in real world environment with strict time constraints.
- Enable product features with **petabytes** level data by Spark, AWS RDS and Airflow.

Education

National Taiwan University, Taiwan, 2009 - 2011

M.S., Department of Computer Science and Information Engineering

National Chiao Tung University, Taiwan, 2005 - 2009

B.S., Department of Computer Science