Yongchan Hong

yongchana@gmail.com | 840-205-6136 | chanhong.xyz | github.com/yongchand

EDUCATION

Harvey Mudd College, Claremont, CA — Class of 2024[1]
B.S. in Computer Science — Overall GPA 3.65, Major GPA 3.9

Coursework: Discrete Mathematics, Differential Eqns/Linear Alg, Biostatistics, Principles of Computer Science, Data Structures & Program Development, Computability & Logic, Algorithms, Computer Science Clinic, Machine Learning

SKILLS

Languages: Python, SQL(BigQuery, MySQL, PostgreSQL), Rust, R, Javascript, Java
Infrastructures: Kubernetes (**Certified Kubernetes Administrator** Holder), Helm, Docker, Terraform, Snowflake, Prometheus, Spark, Tableau, Airflow, Apache Superset
Tools/Frameworks: Git, Next.js, React.js, Linux, Tensorflow, Flask, FastAPI

EXPERIENCES

Data Engineer Intern, Krust Universe, Seoul, Korea — June 2022 - January 2023

- Constructed an overall data pipeline including **EKS**, **Airflow**, **BigQuery** and **Superset/Redash** which provided multiple insights to the team such as total # of transactions, token balances and popular NFT/FTs.
- Generated Klaytn **BigQuery Public Dataset** that requires a backfill of 3+ years of blockchain data; Utilized a load balancing with 15 archive nodes, and performed backfill for 2 weeks
- Created **Open Source** klaytn-etl that enables ETL jobs for Klaytn blockchain data using CLI

Data Platform Engineer Intern, Krafton PUBG, Seoul, Korea — January 2021 - June 2022

- Contributed and utilized Big Data IDE, **Querybook**, as a query platform for the server, data analytics, design team (more than 50+ users created 200+ documents); Created a **Helm Chart** of the Querybook main repository
- Optimized an Apache Superset by customizing **Flask** and **React** codes and presented at 2021 Krafton Developers Connect; Increased possible concurrent users from <20 to 300+
- Created data analytics tool *Rivendell* using **Chakra UI**, **React** and **Flask**; Provided heatmap, custom match data extraction, user match history information which led to data extraction time from 30 min to 1 min
- Built *Crash Report Data Pipeline* using **App Center API**, **Azure Blob Storage & Function**, **AWS S3**, **Lambda**, and **ElasticSearch & Kibana** to provide an organized view of crash incidents to the server team
- Created a Python package using **Spark**, **Pandas** and **Matplotlib** to generate various heatmaps on actual game maps that serve as the basis for ingame design adjustments; Shared during PUBG New State Media Showcase

Software Engineer Intern, Applied Materials BIM team, Santa Clara, CA — May 2019 - August 2019

- Developed simple web pages (e.g., Visa information, **Tableau** redirection, CIF data) that connect to databases based on **PHP**, **HTML/CSS, Javascript** and **SQL Server**; CIF data page was directly used by the VP

PROJECTS/ACTIVITIES

AI/ML in Predicting Protein Properties, Biovia, San Diego, CA — September 2019 - May 2020

- Worked on antibody developability prediction by training machine learning models to predict the Developability Index of antibodies from protein sequence
- Extracted physicochemical features and learned embedding features using **word2vec**; Trained 7 models including **SVM**, **Random Forest**, **MLP** and evaluated the result using **heatmap** and **pareto optimization**

Google Developers Machine Learning Bootcamp, Google Korea — October 2020 - January 2021

- Took Coursera "Deep Learning" and "DeepLearning.AI Tensorflow Developer" coursework and participated in multiple machine learning workshops with startups. Achieved *TensorFlow Developer Certificate*

[1] Worked in engineer roles at companies while serving in the Korean Military (2021-2023)