CHOENDEN KYIRONG

ckyirong@sfu.ca | https://ckyi.me | *Last Updated: October 20, 2022*

EDUCATION

Simon Fraser University   Vancouver, British Columbia
B.Sc in Computer Science & B.Sc in Statistics   *Jan 2018 – Jan 2023*

University of Toronto   Toronto, Ontario
Area of Study: Jazz Performance and Theory   *Sept 2014 – April 2016*

EXPERIENCE

Machine Learning Engineer   Current
BetterCart Analytics: Data Platform Team   *Remote, Canada*
- Designed & deployed the orchestration layer using Prefect onto IBM Cloud in order to automate and productionize Data Engineering & end-to-end ML Pipelines
- Developed a batch NER pipeline using spaCy to separate textual product data into their appropriate attributes & used parallelism to decrease overall runtime of the pipeline by 75% using Dask
- Currently working on the development of a NLP content-based recommendation system

Data Engineering Intern   January 2022 – August 2022
Samsung Research Canada (R&D Institute): Knox Analytics   *Vancouver, B.C.*
- Worked on the automation of a 8hr/week business task by developing an ETL Pipeline in airflow which included extracting and transforming data from S3, creating and managing infrastructure using terraform, and building postgres schemas that are exposed for business to ingest in dashboards
- Gained first hand experience and working knowledge of best practices when working with production level resources while writing and editing queries to production level databases based on specific insights and business needs
- Reduced development time of specific processes by containerizing and streamlining workflows. Designed and built the CI/CD pipeline that automates tests, builds, and deployment of the containers
- Reduced data storage costs by writing python scripts to analyze and report data usage in S3 in order to create Data Retention Policy's.

Data Scientist Intern   May 2021 – August 2021
WorkSafeBC: Innovation Team   *Vancouver, B.C.*
- Improved the speed of document classification by automating the identification and categorization of salient terms using spaCy and the MeSH ontology
- Connected and created a PowerBI tool used by business in order to easily leverage and interact with the salient term recognition system.

PROJECTS

Visualizing Melody, Harmony, & Improvisation | *JavaScript, D3.js, MIDI data, pandas*   Fall 2021
- Researched, experimented, and explored how various musical concepts such as harmony, melody, and improvisation can be aided through visualizations
- Created custom visualizations in D3.js with MIDI data such as: connected graphs of chord tones + shared notes, bi-gram frequency arcs, and an interactive viz using sound to explore and better understand improvisation from Miles Davis' "So What"
- Developed visualizations of personal Spotify data using Spotify's API that shows preferences for certain music features, genres, and artists

Hope Health Action | *Node.js, Express.js, Docker, MongoDB, Mongoose, Gitlab, CI/CD*   Spring 2020
- Led back-end dev for MVC design, data modelling, implementation of the REST API, and built the NoSQL schemas
- Automated development workflow by creating the CI/CD pipeline to handle the testing and deployment of the web application

TECHNICAL SKILLS

Languages: Python, SQL, JavaScript, R
Software Engineering & Databases: Prefect, Airflow, pytest, Dask, Terraform, Docker, AWS, PostgreSQL, GH Actions, flask, Node.js, Express.js, D3.js
ML/DS Tools: PyTorch, Pandas, NumPy, Matplotlib/Seaborn, Scikit-learn, spaCy, dplyr