

CHOENDEN KYIRONG

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

EDUCATION

Simon Fraser University

B.Sc in Computer Science & B.Sc in Statistics

Vancouver, British Columbia

Jan 2018 – Jan 2023

University of Toronto

Area of Study: Jazz Performance and Theory

Toronto, Ontario

Sept 2014 – April 2016

EXPERIENCE

Software & NLP Engineer

Current

BetterCart Technologies: Data Science Team

Remote, Canada

- Worked on the development of a NER tool to separate product data into their appropriate attributes
- Implemented a product matching tool using word2vec embeddings with Genism and Scikit-learn
- Currently working on an ETL Pipeline and the deployment of ML Models

Data Engineering Intern

January 2022 – August 2022

Samsung Research Canada: R&D Institute

Vancouver, B.C.

- Led the development of an end-to-end service that includes an ETL Pipeline in airflow which preprocessed data in S3, creating and managing infrastructure using terraform, and building postgres schemas
- Contributed and produced queries to production level databases based on specific insights and business needs
- Containerized specific development processes that leads to quicker and streamlined workflows and developed the CI/CD pipeline that automates tests, builds, and deployment
- Analyzed and reported data in AWS and created Data Retention Policy's in order to reduce costs
- Worked on the migration of Airflow versions

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

Toxic Comment Classification | PyTorch, Keras, Pandas, Numpy, Matplotlib, Jupyter

Fall 2019

- Achieved 90% accuracy for multi-label classification of toxicity level of tweets
- Leveraged pre-trained 200-D GloVe embeddings to use as input into a LSTM model
- Presented poster and project to industry members, professors, and students at the BC Student AI Showcase

TECHNICAL SKILLS

Languages: Python, SQL, JavaScript, R, HTML/CSS

Software Engineering & Databases: Airflow, Terraform, Docker, PostgreSQL, MongoDB, AWS, GitHub Actions, Node.js, Express.js, D3.js

ML/DS Tools: PyTorch, Pandas, NumPy, Matplotlib/Seaborn, Scikit-learn, spaCy, Tidyverse, dplyr