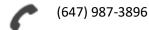
Bill (Yuan Hong) Sun





billyuanhong.sun@mail.utoronto.ca



Education

University of Toronto, Faculty of Applied Science and Engineering

B.A.Sc. in Engineering Science, major in Machine Learning & Artificial Intelligence

Toronto, ON Sep. 2016—Jun. 2021

Skills

- Programming Languages: Python, HTML/CSS, JavaScript, SQL, MATLAB, Java, C, C++
- Data Science: Numpy/Scipy, Pandas, PyTorch, TensorFlow, Scikit-learn, XGBoost; Deep learning and neural networks
- DevOps: Git, AWS, Serverless, Atlassian toolkit (Bitbucket, Jira, Confluence), Docker, Jenkins, Kubernetes
- Frameworks & Libraries: jQuery, ReactJS, NodeJS, Bootstrap, QT, Flask, JWT
- Other: Linux, Bash, cURL, Networking (REST API, sockets, gRPC), databases

Work Experience

Data Science Software Developer Intern

May 2019—Aug. 2020

Toronto, ON

- Nuralogix Corporation
 - Worked in an <u>Agile</u> environment; involved in both development and research work with multiple teams.
 Developed a **full-stack** web application for serving and displaying survey-based machine learning models.
 - Structurally divided into a microservice backend (using AWS Lambda), a Flask frontend, and PostgreSQL database.
 - o Includes a symptoms-based COVID-19 assessment tool, developed from an ensemble machine learning model.
 - o Trained and implemented other machine learning models to predict mental health conditions.
 - Developed an <u>open-source Python</u> API interface to simplify access to the <u>DeepAffex</u> cloud engine, utilizing <u>REST API</u>, <u>WebSocket</u>, and <u>gRPC</u>. Wrote an <u>open-source</u> demo program that shows use cases for the API and SDK.
 - Developed a Python GUI application using QT, along with a Jenkins pipeline that automated data processing.
 - Developed a Serverless web app for uploading and testing videos on newly developed machine learning algorithms.
 - Conducted systematic testing and *quality assurance* for my company's public mobile and web applications.
 - Helped improve existing API software documentation and provided better starter instructions.

Data Science Research Intern—Transportation Modelling Group

University of Toronto Transportation Research Institute

May—Aug. 2017

Toronto, ON

- Analyzed path data from public transit path choice modelling of the Greater Toronto Area.
- Developed fitness functions in Python for path comparison. Improved efficiency by implementing data structures.
- Results were used to improve a machine learning prediction model by over 10% accuracy.

Projects

MusicGenre—Music Genre Classifier Application using Deep Learning

2018

- Multiclass music genre classification utilizing <u>convolutional neural networks</u> (<u>CNNs</u>) and <u>recurrent neural networks</u> (<u>RNNs</u>) from <u>PyTorch</u>. Audio data collection was done using the Spotify API and data scraping methods in Python.
- Used Mel-frequency transform, Fourier transform, and other feature engineering to pre-process raw music audio.
- Achieves > 80% test accuracy in classification. Includes a Flask web interface that samples audio from a YouTube link.

Undergraduate Thesis—Machine Learning Approaches for Rapid Assessment of Anxiety

2020

- Working with Prof. Kang Lee of OISE. Aiming to publish the findings in the CHEST medical journal by April 2021.
- Applied various machine learning techniques to predict the likelihood of anxiety from a questionnaire dataset.
- Selected sets of questions and models that predicted anxiety with over 90% test accuracy; includes a web demo.

Undergraduate Capstone Project—Public Health Ontario

2020

- Applying <u>natural language processing</u> techniques to develop a <u>sentiment analysis</u> model that detects social media posts containing anti-vaccination and other misinformation.
- Developing a web dashboard that scrapes social media posts and updates misinformation statistics in real time.