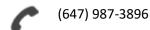
Bill (Yuan Hong) Sun





billyuanhong.sun@mail.utoronto.ca



Education

University of Toronto, Faculty of Applied Science and Engineering

Toronto, ON

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

Sep. 2016—Jun.2021

Skills

- Programming Languages: Python, Java, SQL, C, C++, JavaScript (ReactJS, NodeJS), HTML/CSS, MATLAB
- Data Science & ML: Numpy/Scipy, PyTorch, TensorFlow, Scikit-learn, Pandas; Deep learning and neural networks
- DevOps: Git, Amazon Web Services, Google Cloud, Atlassian toolkit (Bitbucket, Jira, Confluence), Docker, Jenkins
- Other: Linux, Bash, cURL, Flask, networking communication (REST API, sockets, gRPC), Bootstrap, LaTeX, Markdown

Work Experience

Data Science Software Developer, PEY Co-op

Nuralogix Corporation || Toronto, ON || May 2019—Apr. 2020

- Assisting the development of the Anura application, a personal health app using computer vision and deep learning
- Developing a <u>Python</u> API interface to simplify access to the DeepAffex cloud, utilizing REST API, websocket, and gRPC
- Wrote a Python program using DeepAffex SDK and my API client, showing functionalities of the Anura app
- Created an application in <u>Flask</u> (hosted on <u>AWS</u>, using <u>Docker</u>), which takes in data from a user survey, analyzes using a <u>multilayer perceptron (MLP)</u> model, and outputs the likelihood of a mental condition with over 90% accuracy
- Developing a data pipeline in Jenkins to automate and standardize the pre-processing of raw video data

Data Science Research Assistant—Transportation Modelling Group

University of Toronto Transportation Research Institute || Toronto, ON || May—Aug. 2017

- Analyzed path data from public transit path choice modelling of the Greater Toronto Area
- Used <u>Python</u> and its data science libraries to process and compare datasets, by developing different fitness functions
- Results were used to improve a machine learning prediction model by over 20% accuracy
- Improved training time by over 50% through implemented data structures including prefix trees

Projects

MusicGenre—Music Genre Classifier Application using Deep Learning

Nov-Dec. 2018

- Utilizes convolutional neural networks (CNNs) and recurrent neural networks (RNNs) from PyTorch
- Identifies the genre of a piece of music by training on music audio and their genre labels
- Original data collection using the Spotify API and data scraping (with manual data collection)
- Uses Mel-frequency transform, Fourier transform, and other feature engineering to pre-process raw music audio
- Training conducted on TPI service Google Cloud; Achieves over 80% test accuracy
- Includes a web interface (made in Flask and HTML/CSS) that samples audio from a YouTube link

Identifying mobile phone gestures through gyroscope data—PyTorch deep learning project

Oct. 2018

- Utilizes convolutional neural networks to process gyroscope data from mobile phones
- Built and generated own dataset based on data collected in class
- Achieves 85% accuracy in identifying different phone movements

SurroundSound—Music management application for venue hosts (Hack the 6ix submission)

Aug. 2018

- Using React Native with Node.js; utilized MongoDB and Postman API for data storage and retrieval
- Allows users to send music preferences through Spotify API to the venue host and gather most popular choices

Interests