

Anthony Inthavong

anthonyin98@gmail.com | (905) 716 9358 | linkedin.com/in/anthonyinthavong | github.com/anth7310

Education

University of Toronto

Toronto, ON

Honours Bachelor of Science

Sept 2016 - June 2020

- Specialization in Statistical Machine Learning and Data Science

Courses: Machine Learning and Data Mining, Big Data Analysis, Artificial Intelligence, Statistical Inferences, Software Design, Data Structures and Algorithms

Certification

- Google Cloud Platform (GCP) Big Data and Machine Learning Fundamentals ([Coursera](#), 2020)

Technical Skills and Knowledge

- Languages: Python, R, Java, C, SQL, Bash
- Frameworks and Libraries: Pandas, Numpy, BS4, Flask, Django, Regex
- Tools and Technologies: Git, Linux, AWS (EC2, S3, Lambda), GCP, Jupyter notebook, VScode, Excel, JSON

Project Experience

Covid ETL data pipeline and API ([Github](#), [Article](#))

- Developed an ETL (Data Extraction, Transformation, Loading) data pipeline from data collected from John Hopkins University's Github and loaded into a SQL database using Python, Pandas, and Request libraries.
- Used Github's REST API to collect required information so ETL script can be automatically run any time
- Developed extensive REST API using Python and Flask.

Automated Web Scraping and Data Collection ([Github](#): [Rotten Tomatoes](#), [I paid a bribe](#))

- Implemented web scraping programs to collect large amounts of information from various websites.
- Various tools and technologies used depending on the circumstances such as Python, Requests, BS4, Selenium, etc.

Messenger Text Analysis and Dashboard ([Github](#), [Website](#))

- Analytical Dashboard for quick interpretation of your text conversations from Facebook's messenger.
- Unique feature is using IBM Watson's Personality Insights API to summarize text into 5 main personality types.
- Developed backend data processing using Python, Pandas, SQLite, and NLTK (NLP library)

Mental Health in Technology ([Kaggle](#))

- Processed raw survey data spanning several years from *Open Source Mental Illness* into a single SQLite Database.
- Performed extensive ETL processes (manual and automated) using Python, Pandas, and Excel.
- Published dataset to online Data Science community for open sourced research, **43.4k** views and **4981** downloads.

Machine Learning Algorithms and Analysis ([NeuralNets \(CNN\) notebook tutorial](#))

- Implemented various Machine Learning algorithms including but not limited to Linear Regression, Decision Trees/Random Forest, Logistic Regression, Clustering algorithms, and Deep Learning.
- Authored Convolutional Neural Network (CNN) Tutorial using Python, Tensorflow, and Jupyter Notebook

Professional Experience

Data Science Tutor, Freelance ([Reviews on SuperProf](#))

Jan 2019 - Present

- Teaching students fundamentals and best practices of machine learning and data science.
- Various concepts covered such as computer science, mathematics and statistics.
- Highest level of education currently taught is a masters program in Toronto, ON.