

LUONG STEVEN TRUONG

DATA SCIENTIST

EXPERIENCE

Metis · Remote

Data Scientist · Mar. 2021 to June 2021

Completed an accredited 14-week immersive data science and machine learning bootcamp focused on Python, statistics, supervised and unsupervised machine learning, deep learning, data analysis, and visualization techniques.

Drinks and Cocktails Recommendation System Web App

- Built a content-based recommendation system for users to personalize their favorite drinks using Natural Language Processing (NLP) with over 10k data points scraped from multiple sources.
- Leveraged Term Frequency-Inverse Document Frequency (TF-IDF) and Non-negative Matrix Factorization (NMF) for the topic modeling.
- Built and deployed the web app for users to explore their options. The app can be accessed [HERE](#).

Artistic Faces Generator

- Built a neural network model to generate artistic faces using the concept of Deep Convolutional Neural Networks (DCGANs).
- Trained and optimized the model using >200k facial images from the celebA dataset on the Google Cloud Platform in 2 days.
- Built an interactive web app using Streamlit and deployed it to the cloud. It can be accessed [HERE](#).

Customer Complaints Dashboard Web App

- Built an end-to-end data visualization and dashboard to display how customers complain about their financial institutions.
- Used NLTK library to do topic modeling with NLP, Tableau to build the interactive dashboard and embed it into Streamlit.
- Deploy the web app to the cloud for users to explore the analysis themselves. The dashboard can be viewed [HERE](#)

Home Credit Default Risk Classification

- Built a classification model on an imbalance dataset to predict if potential clients are capable of paying their loans if accepted to help Home Credit optimize its process of screening and monitoring its loan applications.
- Used multiple algorithms (XGBoost, Naive-Bayes, Random Forests, Logistic Regression) to compare and choose the best one to optimize.
- XBoost was chosen and the final model has F-2 score of 0.432 and ROC AUC of 0.768.

Car Price Predictor Web App

- Scraped >180k data points from cars.com using BeautifulSoup to build the regression model to predict cars' values.
- Used XGBoost to build and optimize the model. The final model has the R² score of 0.955 and RMSE score of 2788.11 (USD).
- Built an interactive web app to take users' input and estimate the car's values. The app is being deployed on Heroku and can be accessed [HERE](#).

MTA Exploratory Data Analysis

- Scraped the NYC Turnstile dataset to analyze ridership trends post covid-19 for the business use case.
- Visualized and explored data with matplotlib, seaborn, wordcloud, and Foursquare API to seek insights and recommend the best location to open an authentic Vietnamese restaurant in the New York area.

Highline College · Des Moines, WA

Math Tutor · Sept. 2016 to June 2018

- Worked at tutoring center to help students with their homework and test preparation.
- Held over 20 exam review sessions, topics include Calculus, Multivariate Calculus, and Linear Algebra.

Tamarind Tree Restaurant · Seattle, WA

Waitstaff / Bartender · Jan. 2014 to Current

- Manage the medium size restaurant with approximately 10 staffs on duty.
- Highly experienced in customer service and hospitality.
- Experienced in working under high pressure and stressful situation.

EDUCATION

University of Washington - Seattle

Bachelor of Science Mathematics 2020

AWARDS

Highline College · Mathematics Department Certificate Award May 2018

In recognition of outstanding performance in Mathematics

Highline College · The Richard Plagge Mathematics Scholarship May 2017

In recognition of academic success at Highline and accomplishments in Mathematics

CONTACT

✉ tqluong77@gmail.com

☎ 206.565.7237

📍 Seattle, WA, 98125

[in /in/luongtruong77/](https://www.linkedin.com/in/luongtruong77/)

🌐 [luongtruong77](https://luongtruong77.com)

SKILLS

LANGUAGES

Python

SQL

HTML/CSS

MACHINE LEARNING

Linear Regression

Logistic Regression

Naive-Bayes

Random Forests

XGBoost

K-Means

Principal Component Analysis

Natural Language Processing

Convolutional Neural Networks

LIBRARIES AND FRAMEWORKS

Pandas

Numpy

Scikit-learn

XGBoost

Tensorflow

Keras

NLTK

BeautifulSoup

Selenium

DATA VISUALIZATION

Tableau

Matplotlib

Seaborn

Plotly

CLOUD, DATABASE & DEPLOYMENT TOOLS

Streamlit

Heroku

Google Cloud Platform

AWS

GIT/GITHUB

NoSQL(MongoDB)