

## Khanh (Chris) Tran

khanh.tran@simon.rochester.edu | (716) 513-4637 | [linkedin.com/in/chriskhanhtran](https://www.linkedin.com/in/chriskhanhtran) | [chriskhanhtran.github.io](https://github.com/chriskhanhtran)

### EDUCATION

#### **SIMON BUSINESS SCHOOL, University of Rochester** *Master of Science in Business Analytics (STEM-Certified)*

Rochester, NY

May 2020

- Coursework: Core Statistics, Programming for Analytics, Causal and Predictive Analytics, Advanced Predictive Analytics with Python, Text Analytics, Data Management, Big Data, Pricing Analytics

#### **NIAGARA UNIVERSITY**

Niagara University, NY

#### *Bachelor of Business Administration in Accounting*

2019

GPA: 3.99/4.00; Dean's List (all semesters attended)

- Coursework: Business Analytics, Linear Models, Management Information Systems, Econometrics
- Awards: Member of Beta Gamma Sigma Honor Society, Everett Ockerman Award for Academic Excellence

### FEATURED PROJECTS (more details at <https://chriskhanhtran.github.io/>)

#### **Kaggle Competition – Advanced Regression Techniques in Predicting House Prices**

- Performed comprehensive EDA, data cleaning and feature engineering on Ames, Iowa house price data set.
- Built Ridge, Lasso and gradient boosting models to predict house prices, achieving top 6% score in leaderboard.

#### **Predict Breast Cancer with PCA, RF and SVM using Python**

- Performed comprehensive EDA on the Breast Cancer Wisconsin data set.
- Trained Random Forest and Support Vector Machine models to detect breast cancer, achieving 97% accuracy rate.

#### **Detect Spam Messages – Natural Language Processing with Python**

- Tokenized and vectorized text messages using TF-IDF and bag-of-words model.
- Utilized Naïve-Bayes algorithm to classify text messages into normal and spam, achieving 95% accuracy rate.

### EXPERIENCE

#### **TAX TECHNOLOGIES, INC.**

Buffalo, NY

#### **Tax Intern**

Mar. 2019 – July 2019

Provided technical supports to Fortune 500 clients utilizing Tax Series – TTI's flagship product, an all-inclusive SaaS global data collection, tax compliance and provision software.

- Assisted on implementation engagements for new clients, including performing data collection, integration and setup in Tax Series.
- Conducted essential application diagnostics on client financial data, including periodically generating technical reports, maintaining data integrity and monitoring client databases.
- Conducted in-depth research on tax forms and e-file requirements in 32 states and four foreign countries, helping develop annual enhancement release for Tax Series.
- Performed application testing to determine if software works as designed, logged technical reports and collaborated with software engineers to build enhancement update for Tax Series.

#### **BUSINESS ANALYTICS COMPETITION & CONFERENCE**

Manhattan College, New York City, NY

#### **Data Analytics Team Leader**

Feb. 2018 – May 2018

In three-month research and two-day hackathon, led a team of four students to discover insights from NYC and Boston governmental data sets and won runner-up prize for best research poster out of 18 participating colleges.

- Cleaned up (missing data, outlier detection, duplications) and integrated (merge, join, subset) large data sets (6 million records) of governmental spending, contracts and KPI metrics.
- Utilized Python and Tableau to perform exploratory data analysis and visualization on payroll distribution and minority-owned businesses' participation in government contracts.
- Built linear models to determine socioeconomic factors affecting government spending.
- Applied statistical learning techniques to predict government KPI metrics, crime rate and education quality.

#### **ACADEMIC SUCCESS CENTER**

Niagara University, NY

#### **Tutor for Statistics Courses**

Aug. 2018 – Dec. 2018

- Helped nine students understand class materials and build effective study strategies in statistics courses.
- Rated Excellent in all criteria by eight out of nine students.

### SKILLS

**Programming:** Python (Pandas, Scikit-learn, TensorFlow), R, SQL, MATLAB

**Statistical Software and Visualization:** IBM SPSS, Tableau, Python (Seaborn, Matplotlib), Adobe Suite, Excel

**Machine Learning:** Predictive modeling (Regression, Random Forest, SVM), Clustering, Neural Network