

Harrison Cho

harrison_cho@brown.edu

in harrison  hcho1111

[See light theme](#)

EDUCATION

- **Brown University** Providence, Rhode Island
ScM in Data Science - GPA: 4.00 September 2021 – Present
 - Relevant Coursework: Machine Learning Pipelines, Applied Mathematics for Machine Learning, Deep Learning
- **University of North Carolina - Chapel Hill** Chapel Hill, North Carolina
BS: Economics, Minor: Statistics, Highest Distinction and Honors - GPA: 3.88 August 2016 – December 2020
 - Relevant Coursework: Data Analysis Methods, Optimization, Advanced Econometrics, Linear Algebra, Calculus Series
 - Accolades: Phi Beta Kappa, Phillips Ambassador Scholar, Kakehashi Project Representative
 - Thesis: *The Effects of Socioeconomic Characteristics on Ambient Air Pollution and the Decision to Over Pollute*

SKILLS

- **Languages:** Python (scikit-learn, PySpark, TensorFlow, Dask), R (caret, kernlab, shiny), SQL (PostgreSQL & MySQL) Julia
- **Other Tools:** STATA, Excel, Mathematica, \LaTeX
- **In Progress:** NLP in python, web scraping in python, GraphQL, MongoDB, Kafka

RELEVANT EXPERIENCE

- **Watson Institute for International and Public Affairs** Providence, RI
Data Science Research Assistant January 2022 - Present
 - Refining data preprocessing techniques to feature engineer additional spatial and socioeconomic variables related to fatal police encounters in the US
 - Engineering python scripts to web scrape geographic data from the Census Bureau's ACS-5 API given a victim's residency data and fatality location
- **University of North Carolina - Chapel Hill Economics Department** Chapel Hill, NC
Research Assistant December 2020 - April 2021
 - Consolidated literature relating to Monte Carlo simulation, supervised learning techniques, and casual inference conditions in econometric machine learning models to further case study analysis efforts
 - Implemented three machine learning models in scikit-learn and a sequential model in TensorFlow to simulate causal inference techniques for supermarket sales data
- **Jet Aviation Business Jets** Hong Kong
Operations Intern June 2018 - August 2018
 - Authored a process manual outlining unique value streams and risk mitigation protocols in daily operations to accommodate a transitioning senior management team
 - Established client-facing interaction protocols to improve retention of high net worth accounts while maintaining critical quality and safety standards for private pilots

PROJECTS

- **Spatiotemporal Approaches for Classifying Parking Violations** Fall 2021
Python: scikit-learn, Requests, Plotly; API: NYC Geoclient
 - Developed a complete, reproducible ML pipeline via scikit-Learn for classifying 100 unique parking violation categories designated by NYC's Department of Finance. Coupled preexisting geolocation features and NYC's official Geoclient API to engineer granular coordinate data
- **NLP Classification for Dark Web Narcotics Listings** Fall 2020
R: quanteda, caret, dplyr, ggplot2
 - Employed natural language processing techniques to classify clandestine product listings on pre-scraped dark web marketplace data. Researched deep learning techniques to construct a feed-forward neural network, achieving an accuracy score 37% above a standard machine learning model baseline