Harrison Cho

harrison_cho@brown.edu

in harrison hcho1111

See dark theme

</> LaTeX

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 in Economics with Highest Distinction and Honors - GPA: 3.88 — *August 2016 – December 2020*
  - Minors: Business Administration and Statistics
  - Relevant Coursework: Strategic Management, Principles of Marketing, Operations Management, Data Analysis Methods, Advanced Econometrics, Discrete Mathematics, Multivariable Calculus, Linear Algebra
  - Thesis: *The Effects of Socioeconomic Characteristics on Ambient Air Pollution and the Decision to Over Pollute*
  - Awards: Phi Beta Kappa, Phillips Ambassador Scholar, Kakehashi Project Representative

Relevant Experience

- **Watson Institute for International and Public Affairs** — Providence, RI
  Machine Learning Research Assistant — *January 2022 - Present*
  - Refining data provenance techniques to engineer additional spatial and socioeconomic data related to fatal police encounters in the US
  - Assessing statistical approaches analyzing the significance of socioeconomic factors linked to fatal police encounters
- **University of North Carolina - Chapel Hill Economics Department** — Chapel Hill, NC
  Research Assistant — *December 2020 - April 2021*
  - Researched Monte Carlo simulation, supervised learning techniques, and casual inference conditions in econometric machine learning models
  - Implemented machine learning models in scikit-learn and deep learning models in TensorFlow to test the effects of socioeconomic variables on 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 leaner management practices during a transition period
  - Established client-facing interaction protocols to improve retention of high net worth accounts while maintaining critical quality and safety standards

Projects

- **Spatiotemporal Approaches for Classifying Parking Violations** — Fall 2021
  - 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 coordinate data
- **NLP Classification for Dark Web Narcotics Listings** — Fall 2020
  - 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

Libraries, Frameworks, and Technical Skills

- **Python:** xgboost, scikit-learn, pandas, scipy, tensorflow, dask
- **R:** caret, kernlab, dplyr, ggplot, tidyr, shiny, plotly
- **Database Languages:** sql (postgresql & mysql), pyspark, neo4j
- **Other Languages/Modeling Software:** julia, stata, excel, mathematica
- **In Progress:** nlp in python, web scraping in python, graphql, mongodb, kafka