

Harrison Cho

harrison_cho@brown.edu

in [harrison](#)  [hcho1111](#)

</> [hcho1111.github.io](#)

See dark theme

EDUCATION

- **Brown University** Providence, Rhode Island
ScM in Data Science - GPA: 4.00 September 2021 – May 2023
 - Coursework: Applied Machine Learning, Deep Learning, Applied Mathematics for Machine Learning, Statistical Learning, Data Engineering, Applied Econometrics
- **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
 - 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, Keras, PyTorch, OpenCV, PyMC), R (caret, kernlab, shiny), SQL (PostgreSQL & MySQL), Julia, D3.js
- **Other Tools:** Git, Google Cloud Platform, GraphQL, HTML5, STATA, Excel, Mathematica, \LaTeX

RELEVANT EXPERIENCE

- **Ploomber** New York, NY
Software Engineering Intern September 2022 - Present
 - Designed visualization and evaluation tools for distributed ML models in the *sklearn-evaluation* package, adding model comparison, model interpretability, and parallel experiment tracking features
 - Authored a series of technical articles on causal inference modeling in *PyMC*, natural experiment designs for A/B testing, and deploying parallel Bayesian regression discontinuity design experiments in the cloud
- **Brown University Data Science Initiative** Providence, RI
Quantitative Research Intern May 2022 - September 2022
 - Formulated a dynamic cryptocurrency asset pricing model and trading strategy utilizing an ensemble of Machine Learning techniques including, but not limited to: ARIMA + GARCH autoregressive methods, Bayesian Models, CNNs, RNNs (GRUs and LSTMs), and Transformers
 - Before fine tuning practices, our trading strategy achieved a theoretical profit of approximately \$60,000 USD during a bearish, volatile market regime
- **Watson Institute for International and Public Affairs** Providence, RI
Data Science Research Assistant January 2022 - Present
 - Refined data preprocessing techniques to feature engineer geospatial and socioeconomic factors required to assess fatal police encounters in the US
 - Engineered python scripts to automate data collection and entry practices, eliminating manual data validation per fatal encounter (approximately 30% of the time allotted) and reallocating approximately 16.65% of the total research budget

PROJECTS

- **Neural Style Transfer for Algorithmically Generated Camouflage** Spring 2022
Python: TensorFlow, Keras, OpenCV; Cloud Services: Google Cloud Platform
 - Assembled a cloud-based CycleGAN to generate camouflage patterns from 90,000 natural-landscape images. Two additional computer vision models were incorporated with the deep style transfer model for image preprocessing (removal of image watermarks) and feature detection (removal of irrelevant sky-based features)
- **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