Sara Ho

703-568-3229 • sara.a.shiho@gmail.com • linkedin.com/in/Sara-ShiHo • github.com/Sara-ShiHo

EDUCATION

Northwestern University, McCormick School of Engineering Dec 2021
MS in Analytics
Coursework: Deep Learning, Data Mining, Predictive Analytics, Big Data, Data Management, Analytics Value Chains

Virginia Commonwealth University Aug 2017 – Dec 2019
Non-degree STEM Graduate Student
Coursework: Object-Oriented Programming, Advanced Linear Algebra, Advanced Macroeconomics, Regressions

University of Virginia, College of Arts and Sciences May 2017
BAs with Honors in Applied Statistics and Economics

SKILLS

Programming: Python, Java, Go, R, SQL, HTML, CSS, JavaScript, Matlab, Stata, SAS, VBA
Tools and Frameworks: Pandas, PySpark, Sci-kit Learn, TensorFlow, Hadoop, Hive, Flask, Tableau, D3.js, Git, AWS

EXPERIENCE

Data Science Consultant, Intern | ZealStrat, LLC Jun 2021 – Present
- Developing a data-driven efficiency monitoring and risk evaluation system for managing complex R&D projects.

Senior Tutor | Trilogy Education Jun 2019 – Present
- Led over 500 one-on-one online sessions teaching programming and data analytics to over 80 students.
- Maintained a positive and encouraging learning environment; averaged a 4.9 out of 5 on student evaluations.

Data Science Consultant | Unifyd Insights Dec 2020 – Jun 2021
- Implemented multiple deep learning models on messy product data with a focus on testing whether multimodal transformer models can improve upon traditional text-only models. Summarized research in a 2500-word white paper.

Data Science Consultant | United Way Oct 2020 – Jun 2021
- Developed an auto-generated scorecard highlighting constituents' urgent needs to local and state policymakers.
- Analyzed spatial data and time-series trends to highlight special needs during the COVID-19 pandemic.
- Implemented a gradient-boosted tree to identify important predictors of unmet needs.

Data Science Consultant | Feinberg School of Medicine Sep 2020 – Dec 2020
- Digitized hand-drawn scanned images using computer vision denoising and alignment techniques.
- Used results to create informative visualizations of patient pain development.
- Published the summarized results in *The Journal of Pain*.

Research Associate, Full Time | Federal Reserve Bank of Richmond Jun 2017 – Dec 2019
Richmond, VA
- Used Stata, Python, and R to conduct empirical research for publication in academic journals.
- Analyzed large panel data with over 600 million observations on credit default swap trades which provided evidence to support a new theory of financial intermediation in decentralized markets.
- Co-authored a paper on discount window lending to financial institutions in need of short term credit; extracted and analyzed financial data to build case studies on frequent borrowers and the trajectory of their financial health.
- Built and maintained a searchable database of international macroeconomics researchers and their publications.

OPEN SOURCE PROJECTS

WikiNews – Web application providing context to daily headlines
- Built a pipeline to query daily data from multiple API sources using cron and Python.
- Designed an encoding-based recommender system to match news headlines with Wikipedia content.
- Built back end on AWS EC2 and RDS. Built front end using Bootstrap served through a Flask app.
- Implemented version control, logging, metric calculations, unit testing, and Docker containers.

Reviewer | Journal of Open Source Software
Waskom, M. L., (2021). seaborn: statistical data visualization. Journal of Open Source Software, 6(60), 3021.
Mäkelä et al., (2021). Finnish Media Scrapers. Journal of Open Source Software, 6(62), 3400.