ROY N. **WU**

M.S.E. IN COMPUTER AND INFORMATION SCIENCE & DATA SCIENCE

https://roynwu.github.io

+1 (973) 901-1028
wuroy@seas.upenn.edu
linkedin.com/in/roynwu
github.com/roynwu

EDUCATION

Philadelphia, PA — **University of Pennsylvania**
Expected May 2022 — *Master of Science in Engineering, Data Science*
Expected May 2022 — *Master of Science in Engineering, Computer and Information Science*
- **Courses:** Algorithms & Computation, Big Data Analytics, Machine Learning, Deep Learning, Artificial Intelligence, Computer Vision, Computational Linguistics, Time Series Forecasting, Software Foundations
- **Leadership:** Data Science General Assembly, Penn Data Science Group Board

New York, NY — **New York University**
Aug 2013 - May 2017 — *Bachelor of Arts, Mathematics*
Aug 2013 - May 2017 — *Bachelor of Arts, Economics*
- **Courses:** Linear Algebra, Probability, Statistics, Econometrics, Differential Equations, Numerical Analysis
- **Honors and Awards:** Dean's List, University Honors Scholar (Founders' Day Award)

SKILLS

Programming — Python (PyTorch · Scikit-Learn · NumPy · Pandas · Matplotlib · NetworkX) · C++/C · Scala · Java · R
Technologies — SQL · Apache Spark · AWS · Azure Databricks · Git · Docker · Power BI · LaTeX
Others — Actuarial Exams: SOA Exam P/CAS Exam 1 · Languages: Mandarin Chinese · Video: Final Cut Pro, HitFilm

EXPERIENCE

Chicago, IL — **William Blair & Company, L.L.C.**
Jan 2021 - Present — *Data Science Intern*
- Built a data pipeline in Databricks with Python to extract data from different data sources using a combination of Salesforce API and Microsoft Graph API to create data views to be used in Power BI
- Currently designing a network graph and using centrality algorithms to identify effective relationships between bankers and clients

Boulder, CO — **The Trade Desk, Inc.**
Jun 2020 - Aug 2020 — *Data Science Intern*
- Developed a new lookalike modeling feature with Spark and Scala in AWS; this long-awaited function allowed users to customize inputs to the ad campaign optimization algorithm
- Analyzed 10^8+ rows of unexplored ad group data in Vertica SQL and prototyped a custom multi-target linear regression model to support recommender system with Python; this began as an ambiguous project that I drove to completion and launched company-wide, resulting in 74% increase in ad group performance

New York, NY — **Vitech Systems Group, Inc.**
Jun 2018 - Jul 2019 — *Solutions Analyst*
Jun 2017 - May 2018 — *Associate Solutions Analyst*
- Collaborated with tech leads to design extensive workflow processes and write new software design documents, advanced functional specifications, and enhancement contracts
- Taught myself PL/SQL and worked with engineers to write and maintain SQL scripts for extracting reoccurring test cases, cutting down regression testing time by 25%

New York, NY — **Haver Analytics, Inc.**
Sep 2016 - May 2017 — *Economic Research Intern*
- Compiled, managed, and quality-assured databases covering macroeconomic indicators of Asia-Pacific markets and delivered data upon each release to clients with DOS
- Added new time-series data to databases and automated data extraction and formatting with DOS

PROJECTS

Academic Research — **Generating News Headlines Using Deep Learning Abstractive Summarization [paper] [code] [blog]**
Jan 2020 - May 2020
- Trained a Seq2Seq network in Pytorch on 200,000 news articles to generate summary headlines
- Improved the network with attention mechanism, teacher forcing, and a pointer-generator; achieved BLEU score of 0.075, outperforming results from main reference paper from 2015 by 0.065

Academic Research — **Artificial Anime Character Design Using Generative Adversarial Networks [paper] [code] [blog]**
Jan 2020 - May 2020
- Implemented generative adversarial networks in Pytorch with improved training techniques (trained on 21,000 images) to generate artificial anime character designs