Christina Bogdan

christina.bogdan@gmail.com • christaina.github.io • (347) 306-7370

EDUCATION

New York University, New York, NY

Master of Science (MS) in Data Science — Sep 2015 – May 2017

- Courses: Machine Learning, Deep Learning, Statistical and Mathematical Methods for Data Science, Fundamental Algorithms, Natural Language Understanding with Deep Learning, Big Data, Inference and Representation, Monte Carlo Methods, Causal Inference, Advanced Machine Learning

Bachelor of Arts (BA) in Mathematics — Sep 2013 – May 2015

- Courses: Honors Linear Algebra (Graduate), Basic Probability (Graduate), Data Structures, Mathematical Modeling, Mathematical Statistics, Theory of Probability

PROFESSIONAL EXPERIENCE

The Climate Corporation, San Francisco, CA

Senior Data Scientist — Mar 2018 – Present

- Implemented production pipeline for seed placement product using Pyspark and Airflow
- Worked with researchers to refine evaluation and other components of seed placement model to prepare for production
- Created internal Python package for model evaluation
- Hold office hours for team's data science platform. Answer questions from researchers on data science best practices
- Built POC of model to flag demo fields using clustering
- Advised researchers in multi-company collaboration on deep learning methods to predict yield during growing season

The Climate Corporation, San Francisco, CA

Data Scientist — Jul 2017 – Mar 2018

- Created compressed representations of weather time series data using autoencoders for use as features
- Implement pipeline to train deep learning models at scale with Spark
- Wrote post for tech blog on process of operationalizing distributed deep learning on our Spark platform

NYU Department of Environmental Studies, New York, NY

Research Assistant — Oct 2016 – Jun 2017

- Applied convolutional neural networks to identify wildlife trade from online retailers using image and text data
- Improved models by creating Word2Vec embeddings using large amounts of unlabeled data
- Built end to end model to scrape and clean data, classify item properties, and store results using Python, TensorFlow
- Set up and maintain a MongoDB database of potential wildlife trade and model results

Knewton, New York, NY

Data Science Intern — Jun 2016 – Sep 2016

- Designed and implemented methods for diagnosing skill levels of new students with bandit approaches in Java
- Formulated metrics to evaluate these policies and ran simulations of students to assess different approaches
- Investigated quality of student simulations and suggested improvements
- Implemented Bayesian optimization algorithm in Python for quickly optimizing parameters of recommender to automate QA using simulations

The Weather Company, New York, NY

Part-Time Data Analyst — May 2014 – May 2016

- Automated daily ETL process to integrate data from multiple sources and load in Amazon Redshift using Python
- Held a series of seminars instructing team members on using Python for data analysis
- Created dashboards in Tableau and QlikView summarizing advertising trends

PROJECTS

Visualizing Twitter Networks of the 2014 Ukraine Crisis, MS in Data Science, NYU — Fall 2016

- Created general heuristic to restructure and reduce large retweet network while preserving information
- Designed interactive visualization in D3.js tracking crisis over 95 days
- Won 'Best Data Visualization' at the NYU Center for Data Science Academy Awards

Question-Answering Using Deep Learning, MS in Data Science, NYU — Fall 2016

- Implemented model (https://arxiv.org/abs/1606.02858) for question answering in TensorFlow
- Experimented with model architecture using different methods of embedding and attention

LANGUAGES & TOOLS

Python (advanced), SQL (advanced), Java (intermediate), Javascript (intermediate), R (intermediate), MATLAB (novice), Lua (novice), Scala (novice), LaTeX, Vim, AWS, Git, Pandas, TensorFlow, Keras, scikit-learn, Torch, D3.js, MongoDB, Spark, YARN, Docker, Apache Airflow