

Dylan Labatt Randle

dylanrandle.github.io

Email : dylanrandle@g.harvard.edu

Mobile : +1-647-641-1994

EDUCATION

- **Harvard University, School of Engineering & Applied Sciences** Cambridge, MA
Master of Science in Data Science; GPA: 4.0
Sep 2018 – Present
 - **Awards:** Applied Computation Scholarship
- **University of California at Berkeley, College of Engineering** Berkeley, CA
Bachelor of Science in Industrial Engineering & Operations Research; GPA: 3.9
Sep 2012 – May 2016
 - **Awards:** High Honors, Phi Beta Kappa, Dean's List (2012-2016)

RELEVANT EXPERIENCE

- **Harvard University** Cambridge, MA
Graduate Researcher
Nov 2018 – Present
 - **Deep Learning in Physics:** Researched and developed methods for solving Navier-Stokes equations using unsupervised neural networks. Researching methods to train GANs to solve differential equations (e.g. harmonic oscillator, Burgers' equation) in an unsupervised fashion.
- **Amazon Robotics** Boston, MA
Data Science Intern
Jun 2019 - Aug 2019
 - **Data Engineering:** Built automated, scalable data pipeline for big data queries, data merging, data cleaning, and data transformation. Tech stack: Python (Pandas, Dask, Boto3), Apache Spark, AWS EMR/Athena
 - **Data Science:** Developed machine learning library for proprietary internal product. Included data filtering, feature selection, linear models, random forests, gradient boosted trees, and feed-forward neural networks. Leveraged interpretability algorithms (e.g. feature importance, ALE, SHAP) for model explanations. Tech stack: Python, Pandas, NumPy, scikit-learn, Keras, XGBoost, AWS SageMaker
 - **Data Product:** Developed API for easy development of ML models. Built front-end UI for non-technical users. Demonstrated results to stakeholders. Tech stack: Python, Jupyter, IPython Widgets
- **Hubdoc** Toronto, Canada
Lead Data Scientist
Jan 2017 – Jul 2018
 - **Production Deep Learning:** Developed and deployed production deep learning system using LSTMs & CNNs for entity extraction and text classification of financial documents. Models trained on over 10 TB of text and image documents. Extraction time reduced from 24 hours to 5 seconds. Cost savings estimated at \$2MM/year. Tech stack: Python, Keras, Tensorflow-Serving, PostgreSQL, AWS EC2
 - **Data Science:** Conducted business and engineering analyses: e.g. prediction of labor requirements and anomaly detection of web scrapers. Wrote reports and built data visualizations for company intranet in D3.js
 - **Leadership:** Regularly presented results and recommendations to C-suite. Integral in crafting team strategy and roadmap. Involved in fundraising and presentations to investors. Delivered machine learning lecture to 60+ people
- **Taylor Statten Camps** Algonquin Provincial Park, Canada
Canoe Trip Guide
Summers 2015, 2016
 - **Canoe Trips:** Led 36- and 50-day canoe trips through remote Canadian wilderness. Responsible for groups of 7 teenage boys. Responsible for planning, safety, and navigation
 - **Camp Maintenance:** Built a new dock; renovated and painted cabins. Leader of roofing crew
- **Bank of Montreal, Capital Markets** Toronto, Canada
Financial Products Analyst
Summer 2014
 - **Fixed Income:** Conducted analyses of various debt products (swaps, swaptions, ABS, MBS). Wrote custom C# algorithm to analyze relationship between delta-hedging frequency and returns for Canadian swaptions; found possible trading opportunities
 - **Sales & Trading:** Compiled daily summaries of trading activity. Reviewed and analyzed sales product pitches. Supported both sales and trading with various data analyses

RELEVANT PROJECTS

(Titles link to project websites.)

- **Twitter Troll Classification:** Project achieving 96% accuracy classifying Twitter trolls using tweets scraped from accounts indicted for meddling in the 2016 U.S. elections
- **Automatic Differentiation:** Python package implementing automatic differentiation, supporting both forward and reverse modes; stochastic gradient descent and Adam optimizers implemented as example use-case
- **Bayesian GANs:** Paper review, implementation, and demo of Bayesian generative adversarial networks (GANs)
- **Tensorflow on Spark:** Training neural networks on a 1.5 TB dataset with Tensorflow on a Spark/Hadoop cluster with AWS Elastic Map Reduce
- **Microbiome Dynamics:** Modeling Granger causality with causal-LSTM model of high-dimensional experimental microbiome time-series data from mice
- **Safe Autonomous Vehicles:** Critical thinking project demonstrating methods (federated learning, differential privacy, secure multi-party computation) and evaluating policies

PROGRAMMING SKILLS

Languages: Python (NumPy, Pandas, scikit-learn, PyTorch, Keras, PyMC3, Boto3), SQL, C

Technologies: AWS (EC2, EMR, S3, Athena), Hadoop, Spark, OpenMP, OpenACC, MPI, Git, LaTeX, Markdown