

MATIAS BERRETTA

241A Madison St, Apt 3, Brooklyn, NY 11216

440-935-9201 — matiasberretta93@gmail.com — <https://mberrett.github.io/>

EXPERIENCE

Data Scientist

June 2019 — Present

Healthfirst

- Saved 150k in third party vendor contracts by bringing churn modeling for Medicare voluntary disenrollment in-house (AUC + 20%)
- Key player in HF rising to a 4-star Medicare rating by reducing Medicare voluntary disenrollment through risk stratified outreach
- Developed first Deep Learning use case at HF by implementing Recurrent Neural Network for Preventable Inpatient Admissions
- Developed NLP Topic Modeling Module to help Contact Center Operations understand why members call in order to drive strategy
- DS Team lead for adopting AWS services (EC2, S3, etc.), GitHub (version control, project doc, etc.), and new tech in general
- Reduced code run time overall by 50% by migrating new and existing projects to AWS EC2 virtual machine
- Reduced time needed for feature-engineering by 75% by teaching team GitHub and cataloging all known features into repositories
- Optimized Machine Learning workflow (robust + resilient) with the first implementation of Spark within the Data Science team
- Set up pipeline (EC2, S3, PySpark, Redshift) for ingesting and integrating third-party (Facebook) data: 96 million rows in 5 min
- Org-wide Data Science ambassador, presented on Deep Learning, GitHub to various high level executives (CTO, CSO, CIO)

Data Analytics Graduate Assistant

August 2018 — May 2019

Fordham University and Harvard Medical School

- Built customized predictive models to help doctors at Harvard prescribe the right treatment for patients with multiple sclerosis
- Developed a new, intelligent method of imputation for missing values in multivariate time series called dynamic fill
- Contributed a new variation on GRU called GRU-DF, which incorporates dynamic fill into GRU's training operation

Technical Solutions Consultant Intern

June 2018 — August 2018

Google Inc. (G-Tech)

- Developed ticket-theming tool to quickly surface trends for G-Tech services such as Ad Words using Natural Language Processing
- Obviated ticket misclassification issues and increased processing speed and quality for Product Operations Managers
- Presented project to Vice President of G-tech

Data Analytics Graduate Assistant

August 2017 — May 2018

Wireless Sensor Data Mining (WISDM) Lab at Fordham University

- Advised software engineers in the lab on implementation of machine learning and natural language processing models
- Web-scraped and parsed through the dark net to collect training data for terroristic threat classifier (approved by the FBI)

EDUCATION

M.S. in Data Analytics

Fordham University GSAS, New York, New York

Expected Graduation May 23, 2019

GPA: 3.96

B.A. in Psychology and Creative Writing

Oberlin College, Oberlin, Ohio

Graduated May 23, 2016

GPA: 3.69

RESEARCH

GRU-DF: A Temporal Model with Dynamic Imputation for Missing Target Values in Longitudinal Patient Data

Accepted at ICHI 2020: IEEE International Conference on Healthcare Informatics, one of the top conferences in biomedical and medical informatics, with a ~30% acceptance rate and a double-blind peer-review process

The Earth is our Home: Systemic Metaphors to Redefine our Relationship with Nature

Published: 23 February 2017. Accepted at BECC 2017: Behavior, Energy and Climate Change.

SKILLS

PROGRAMMING LANGUAGES AND TECHNOLOGIES

Python (Pandas, NumPy, TensorFlow, Keras, PySpark, Sklearn, Matplotlib, Seaborn), Jupyter Notebook, R (ggplot2), R-Studio, PostgreSQL, AWS Cloud Services, Linux, Git, Spark, Linux, Tableau, Excel

LANGUAGES

Native Speaker in English, Spanish, and Portuguese. Conversational Proficiency in French. Basic Proficiency in Japanese.