

Jeremy Miller | Data Scientist

415.902.6812 | jeremymiller00@gmail.com | San Francisco CA

Linkedin: linkedin.com/in/jeremymiller00/ | Github: jeremymiller00

SKILLS

Machine Learning | Prediction and Classification Algorithms | Feature Engineering | Data Visualization and Analysis | Data Cleaning and Transformation | A/B Testing | NLP | Deep Learning | Frequentist and Bayesian Statistical Inference | Python | Apache Spark | SQL | Git | Scala | AWS Cloud Computing

EXPERIENCE

Data Scientist, Clarivate Analytics

2019-current

Research, develop, validate, deliver, and monitor Machine Learning solutions to improve user experience with Clarivate's data and services. Machine learning solutions in production include:

- Delivered custom Spark application to auto-generate document cluster labels for clusters of scholarly documents ranging in size from 10 to 10,000. Scored 92% precision on test dataset. Worked with engineering and QA teams to integrate application into new product.
- Increased recall@10 by 15% for ~15,000 class Natural Language classification problem by delivering updated Deep Learning model with Tensorflow. Refactored Tensorflow code to enhance forward compatibility; implemented automated workflow in Databricks for quarterly data updates and model retraining.
- Collaborated with team members to produce custom Java library for bespoke term extraction use case, saving the company ~\$96k per year and allowing a ~\$15mil product to continue operating smoothly.
- Provided key algorithm in product with revenue ~100k per year, a Logistic Regression model for Natural Language document classification.
- More details: <https://github.com/jeremymiller00/ProjectList>

Data Scientist in Residence, Galvanize

2019

Instructed and supported Data Science Immersive students in Python programming, statistics, and machine learning.

- Assisted in teaching Data Science curriculum; worked with students in groups and individually to develop their data science skills and knowledge.
- Facilitated assessments, warmup sessions, team case studies, and individual projects.
- Contributed to development and refinement of Data Science curriculum.
- Continued ongoing development of data science practice in conjunction with core Galvanize faculty.

Academic Coordinator, FIDM San Francisco

2007-2018

Managed and lead campus tutoring and academic support programs; coordinated Bachelor of Science in Business Management degree program.

- Optimized tutor scheduling through data collection and analysis, resulting in increased visits to learning center and greater student engagement.
- Developed alternative course scheduling plans for Business Management student to allow cohorts to start at any time of year.
- Worked with leadership team to develop and implement Student Success Seminar, resulting in increased performance from academically struggling students.
- Solved numerous individual academic problems with students by carefully balancing their needs with institutional requirements and limitations.

EDUCATION

Data Science Immersive, Galvanize, San Francisco

2018

Master of Arts in Art History, San Francisco State University, College of Creative Arts

2007

Bachelor of Fine Arts in Photography, San Francisco Art Institute

1999

SERVICE

National Laboratory for Education Transformation, CompareK12 Data Systems Advisor

Volunteer Tutor, 826 Valencia, Tutor for under-resourced students ages 6-18

Volunteer, San Francisco-Marin Food Bank Deliver food to families in need

Volunteer, Code for America, SF Brigade, Data Scientist