

JEROME RUFIN

DATA SCIENTIST

CONTACT

✉ jeromerufin@gmail.com
☎ 925-326-8078
📍 New York, NY 11211

in jerome-rufin-a1911275/
🔗 /Jeromerufin

SKILLS

TOOLS: Tableau, Zeppelin, JIRA, AWS, Command Line, github, Jupyter, Excel
TECHNOLOGIES: HTML/CSS, Flask, Javascript, Python, SQL
DATABASES: PostgreSQL, MySQL, BigQuery, Oracle
SPOKEN LANGUAGES: English - Native, Tagalog - Intermediate
PACKAGES: Scikit-Learn, NumPy, Pandas, NLTK, Selenium, BeautifulSoup, Seaborn, Matplotlib, Vader, spaCy, Keras, XGBoost
STATISTICS/ML: Linear/Logistic Regression, Random Forest/Gradient Boosting, NLP (LSA/NMF/Sentiment Analysis), Neural Nets (Auto encoders)

EXPERIENCE

General Assembly

New York, NY

Data Science & Analytics Instructor Associate · Jan. 2020 to Current

- Teaching Data Science and Data Analytics to industry professionals through GA's Enterprise Program (client: Booz Allen Hamilton).
- Deep Diving into SQL with PostgreSQL, Visualization with Tableau, data/statistical analysis with Python, and Machine Learning (i.e. classification, regression, and decision trees)

Metis

New York, NY

Data Scientist · Sept. 2019 to Dec. 2019

NOTE: please see project section

- Metis's 12-week Data Science Bootcamp is a full-time, immersive program revolved around modeling techniques, statistical analysis, and machine/deep learning

Twitter

San Francisco, CA

Data Analyst · June 2018 to Sept. 2019

- Analyzed Twitter's platform health using tools including SQL/Python to create heuristics that remediated over 5M spam and abuse accounts
- Led the initiative to develop metrics, using Tableau and Zeppelin, that reduced operational workload by 150%
- Created a specialized ML focused team called GRECO (please see project section)

Visa Inc.

San Francisco, CA

Data Analyst (promotion) | Business Dev. Associate · July 2016 to May 2018

- Performed Data Analysis with SQL/Tableau for 200+ Visa partners mounting to over \$3MM in net yearly revenue
- Handled business strategy and managed marketing campaigns for 25 accounts (\$5MM in net yearly revenue)

PROJECTS

GRECO - Global Rule Engine for Customer Operations (Twitter)

- Developed a rule engine to remediate users violating Twitter's spam policies
- Discovered features with SQL and classified classes with tree-based methods (i.e Random Forest)
- Reduced agent backlog by > 10% and was estimated to save over \$2M/year in agent case handling time.

Proactive Abuse Detection (Twitter)

- Increased proactive abuse detection for the harmful spread of private information by 38%
- Used SQL to identify patterns/features to develop heuristics for detection
- Created automated bots that captured ~8k offenders with ~80-90% precision.

Recommendation System - Automated Playlist Generator (Metis)

- Created a CNN/RNN powered engine to recommend playlists based off of emotion and music preference
- Utilized Librosa and Auto-encoders to calculate cosine similarities in order to make the recommendations

Classification on eSports (Metis)

- Calculated winning probabilities (AUC: 0.85) of eSport matches using Logistic Regression and Gradient Boosting
- Web-scraped, using Selenium, to extract data from multiple sources

Fake News NLP Investigation (Metis)

- Pre-processed tweets by the US government by developing a Spacy/NLTK pipeline
- Performed Topic Modeling using TF-IDF/NMF and sentiment analysis using VADER
- Trained a neural net (GPT-2) and applied heuristics to better understand how to fight fake news

Predicting NBA scores with Linear Regression (Metis)

- Improved linear regression model's R^2 through polynomial transformations and feature engineering
- Predicted winner of a game 69.6% of the time (ESPN's algorithm has a predicting power of 70.3%)

EDUCATION

University of California, Berkeley

BS Industrial Engineering and Operations Research 2016

General Assembly · Jan. 2018 to Feb. 2018

Data Analytics

University of Chicago Graham School · Jan. 2019 to Feb. 2019

Project Management