

Rushil Jagat Sheth

sheth.rushil@gmail.com | rushilsheth.com | (314) - 766 – 3990 | San Francisco

Education

University of San Francisco San Francisco, CA
Master of Science in Data Science

June 2020

University of California – Berkeley Berkeley, CA
Bachelor of Arts in Applied Mathematics
Bachelor of Arts in Statistics

May 2017

Work Experience

Machine Learning Engineer Lead

February 2022 – Present

Pathlight

San Francisco, CA

- Developed and implemented generative algorithms via LLMs from 0→1, ensuring autonomous, precise, streamlined insights generation through cutting edge research, RLHF, LangChain, and vector databases
- Directed technical projects autonomously, meticulously scoping requirements and sequencing work for the engineering team, and serving as the primary LLM and ML expert advisor to the CTO and team
- Contributed to the creation of an external service using AWS Lambda for large-scale machine learning models. This effort focused on offloading computational tasks to enhance web app performance, and involved the integration of pgvector as the embedding database for optimal LLM accuracy and efficiency
- Hosted and fine-tuned open source LLMs for cost reduction and flexibility on task routing
- Spearheaded the integration of machine learning techniques into product features, significantly contributing to the company's innovation pipeline and competitive edge in conversational analytics [\[blog post\]](#)
- Collaborated with teams across departments leveraging dbt and Hex for advanced analytics and automated the processing of customer data requests, boosting satisfaction as the sole Data personnel at the company

Data Scientist

July 2020 – February 2022

eHealth

San Francisco, CA

- Identified customers likely to churn through a Random Forest classifier and MLflow for model selection, which resulted in a high AUC and recall. Results were made accessible via an ETL pipeline to Snowflake
- Determined lifetime value (LTV) for customers using a Cox Proportional Hazard model. Delivered results via AWS SageMaker and Lambda to create a REST API
- Created an end-to-end model, pipeline, and monitoring to predict the LTV associated with sending a mailer to an individual via Spark, XGBoost and H2O which led to over 80% LTV with a 50% reduction in cost

Data Science Intern

October 2019 – June 2020

New York Mets

Remote

- Maximized outs using statistical distributions, XGBoost, and K-means to create an infield and outfield defensive shift model in an accessible dashboard hosted on AWS
- Enhanced organization's understanding of fielders' ability with a novel internal metric, plus/minus, created with a Random Forest classifier, cross entropy, and data aggregation in pandas

Business Analyst

September 2018 – July 2019

Pinterest

San Francisco, CA

- Created a Tableau dashboard to influence decision making and improve process efficiency. Decreased the duration of time required to solve outstanding service tickets by 30%

Proficient Technologies: Python, LLMs, PyTorch, ML, MLOps, MLflow, SQL, Next.js, AWS