

Arun Natesan

Austin, TX

Experience

Procore Technologies (2022 – 2025):

Senior Machine Learning Engineer

- Built the core risk model used in underwriting material financing transactions. The model reduced loss rates by over 80%.
- Contributed to Automatic Data Capture work: extracting data from construction documents using OCR, and then running classification, NLP, and generative models on the data to create accurate, useful labeled data for a multitude of downstream use cases.
- Developed a tree-based Operational Excellence model using behavioral metrics to predict project outcomes. Leveraged SHAP values to generate normalized, user-specific recommendations integrated into Procore's LLM powered Copilot for real-time next best action guidance

Affirm (2021 – 2022):

Senior Data Scientist

- Built the Merchant Lifetime Value model used for marketing and merchant partnerships.
- Led a team that built the Merchant Churn Risk model.
- Ran A/B tests improving the merchant onboarding funnel and post-onboarding experience which led to a 55% increase in merchant activation and a 12% increase in merchant retention.

Earnin (2018 – 2021):

Senior Data Scientist

- Built the Customer Lifetime Value model and underlying data structures used for company-wide finance and growth planning. Also used as an important feeder model for the core risk model.
- Built the marketing spend optimization model as well as the underlying marketing attribution model used to make over \$50M in marketing spend decisions.
- Led a team that built a complex, near real-time data pipeline to enable researchers including Harvard's [Opportunity Insights](#) group, to study the economic impact of COVID-19 and subsequent policy decisions while maintaining strong data anonymity. [NYTimes article](#) [Bloomberg article](#)

Stanford Children's Health | Lucille Packard Children's Hospital (2017):

Operations Research – Data Analyst

- Analyzed data to identify underutilized resources and worked with staff to implement strategies to increase utilization. Our suggestions increased recovery room throughput by 10%.

Stanford Institute for Economic Policy Research (2016 – 2017):

Ballmer Group – Research Assistant

- Worked to create a [data profile](#) of the US government with detailed accounts of spending and key performance indicators (KPIs) for each sector of spending.
- Presented sections of the data to Steve Ballmer and the rest of the team to ensure quality and clarity.

Deloitte Consulting (2013 – 2015):

Rhode Island Unified Health Infrastructure Project (UHIP) – Project Analyst

- Wrote eligibility rules for all social services and assisted with the database migration.
- Interacted with employees of the state of Rhode Island to determine user needs and managed a team of 4 developers to ensure user requirements were met.
- Showcased our system to top-level state partners including the Governor of Rhode Island.
- Led User Acceptance Testing for both phases and supervised a team of 20-25 testers each time.

Education

Stanford University (2012 – 2018):

B.S. – Management Science and Engineering (MS&E)

- Operations Analytics Focus

Skills

- Strong understanding of probability, statistics, and many domains of machine learning (NLP, classification, regression, anomaly detection, clustering, etc.)
- Experienced working with large structured and unstructured datasets in a distributed environment
- Working knowledge of modern ETL, ELT, and data infrastructure practices
- **Software and Programming Languages:** Python, R, SQL, Airflow, Kafka, Rest, MLflow, DBT, AWS (S3, EC2, Lambda), Databricks, LangChain, Frontier LLMs