

Jason Shiverick

Lead Data Scientist and Data Engineer

jason.shiverick@gmail.com
linkedin
415-849-5589

Tech Stack

- Code : Python, Spark, SQL / NoSQL, Go
 - DevOps : Git, Docker, Ansible, Airflow, NGINX
 - AWS : EC2, S3, Glue, Athena, RDS, Lambda, Kinesis, serverless.js
 - Hadoop : HDFS, Hive, Impala
 - Stats and ML : pandas, scipy.stats, numpy, sklearn, lifelines, pymc3, MLlib
-

Experience

Waymo [Contract] *Senior Data Scientist* June 2018 to Present

- Designed and developed the python run feature extractor pipeline. Built on apache beam, flume, borg, python and protobuf, the framework provides users a simple design pattern to extract features from vehicle and sds log data, and then to scale over the whole fleet processing 100's of TB in minutes.
- Developed Metropolis Hastings algorithm in Go for sampling from a Weibull posterior with arbitrary priors.
- Developed Markov chain Monte Carlo python code base for accurately forecasting field failures in complex systems.
- Developed Reliability analytics data pipeline and dashboards for report automation.

Mayfield Robotics [Contract] *Data Engineer, Consultant* March 2018 to June 2018

- Data Warehouse: Designed and implemented analytics data infrastructure using spark via AWS Glue to process robot logs and disparate data sources into AWS Athena optimized parquet files on S3.

Tesla *Software Engineer, Data Scientist, Manager* 2013 to 2018

- Built and Maintained robust back end infrastructure on top of Docker and Ansible. My design made it easy to provision and manage a Spark cluster and various Micro-Services between two people while also providing statistical models and TB scale log analytics.
- Established an analytics workflow leveraging git version control, with jira integrations. Designed the ETL workflow using spark, airflow, jupyter and superset.
- Provided direction on proactive maintenance campaign and prognostics algorithm development using machine learning techniques: *random forest, logistic regression, physics of failure.*
- Established an extensive code base that provides tools to the organization for extracting, transforming, and analyzing field data at scale.
- Developed a modern approach to advanced warranty simulation in Python that can account for competing failure modes in a repairable system under varying use conditions.
- Developed statistical frame work for python: *Weibull analysis, Stress-Strength Convolution, Hypothesis testing, Best fit solver, generalized distribution framework, newton-raphson solver, ranking methods, mttf*

Ingersoll Rand *Reliability Engineer* 2011 to 2013

Medtronic INC. *Product Performance Specialist* 2010 to 2011

Boeing Corporation *Systems Engineer* 2008 to 2009

Education

Graduate Course Work (Reliability Engineering) 2012 to 2013 University of Maryland (online) College Park, Maryland *ENRE 602: Reliability Analysis ENRE 655: Advanced Methods in Reliability Modeling*

Invited Talks

PHM Society 2015 automotive panel discussion

ARS 2014 Big Data in Reliability: 1st Place