

# SEPEHR SADIGHPOUR

STAFF DATA ENGINEER / MANAGER

[hi@sepehr.xyz](mailto:hi@sepehr.xyz)

(510) 759-4850

Long Beach, CA

## Profile

I am a machine learning and data engineer with experience leading teams to architect scalable solutions that help healthcare companies innovate more quickly.

## Strengths

### ML DevOps

Creating workflows for data scientists to rapidly train, interpret, productionize, and monitor ML models.

### Data Engineering

Building ETL and ML pipelines using Spark and SparkML. Empowering analysts to validate data quickly and surface insights.

### Research

Deploying solutions in machine learning and statistics to develop tailored business solutions.

## Experience

### Actionium Health (Formerly "SymphonyRM")

Mar 2020 - Present

#### Staff Data Engineer / Manager

Managed and mentored team of 4 data engineers, acting as tech-lead and scrum master. Introduced rigorous standards of validating data pipelines and monitoring system behavior. Helped co-architect an event-driven platform for ingesting data.

### Bodyport

May 2020 - Mar 2020

#### Senior Data Engineer

Accelerated time-to-market by developing dashboard product for visualizing physiological time-series signals from device in real-time. Minimized risk of catastrophic loss by bringing all data-science related infrastructure under code, creating a completely reproducible tech stack. Created developer tools for generating realistic data while protecting patient privacy.

### Clarify Health Solutions

Nov 2017 - Mar 2020

#### Senior Data Science Engineer

Architected ML training and tracking framework capable of parallel training and serving 1M+ models in production in batch and real-time modes. Managed team of developers to implement. Implemented massively distributed model interpretation. Built Spark pipelines for data validation, feature engineering, and model training on 10s of billions of rows of healthcare claims and social determinants of health data.

### Omada Health

Apr 2016 - Jul 2017

#### Data Scientist

Guided product direction via analyses of trends in participant behavior, leading to 6% lift in user engagement. Designed and implemented A/B tests for new features in close collaboration with designers. Managed 4-person data team's research into areas such as chatbots, reinforcement learning, and machine vision.

### BayThrive

Jan 2010 - Apr 2016

#### Owner, Developer, Project Manager

Increased traffic 20x and conversion rates 4x for biotech sales firm through SEO and iterative design. Built development and production infrastructure for early-stage startup using Docker and Ansible. Developed heuristic for financial research firm to estimate degree of self-similarity in time-series data.

### Mepkin Abbey

Jun 2008 - Jan 2010

#### Webmaster, Go-To Guy

Directed development of custom registration system from need-assessment to hiring and project management  
Built website, e-commerce solution, and established processes for inventory management and online sales

---

## Education

### Galvanize University

Oct 2015 - Jan 2016

#### Data Science Fellow

Capstone Project: Built information-retrieval app for discovering scientific articles by semantic similarity using document vectors

---

### Duke University

Class of '08

#### Bachelors of Science, Physics

Capstone Project: Trained neural networks and genetic algorithms to solve energy states of quantum systems

---

## Technical

Modeling: Spark ML, scikit-learn, H2O, Vowpal Wabbit, LIME

NLP: Spacy, gensim, fastText, WordNet, NLTK

ETL: Spark, Airflow, Luigi

Infra: Terraform (AWS EMR, Batch, ECS, S3, Redshift), Docker

Software: Flask/Django, PHP, git, bash, RESTful API design, some JS

Analysis: Pandas, R, SQL (Postgres, Presto)

Experimentation: Experiment design, multi-armed bandits

Viz: Plotly, leaflet.js, seaborn, some D3.js