

# ZOE FARMER

ML, Modeling, & AI Engineer

## Staff Machine Learning Engineer Senior Data Scientist

Square (San Francisco, CA)

04/21–Current

Built, tuned, and deployed machine learning models to support high-value initiatives after extensive work identifying needs with product stakeholders,

- Marketing lead conversion-propensity ranking with LightGBM
- Public web traffic classification using PyTorch+Keras
- Anomaly detection on high-density hardware metrics via Isolation Forests
- Product lifetime-value modeling using linear methods

Established mature data & machine learning ecosystems to support modeling work.

Supported a large product team with a wide array of data science and modeling requests while additionally managing contracted roles to provide additional team bandwidth.

## Data Scientist

Red Dot Storage (Louisville, CO)

04/19–04/21

Built, tuned, and deployed financial models to optimize the core revenue management through dynamic churn prediction.

Supplied leadership with a narrative-driven approach to critical insights into customers, markets, and strategy via statistical modeling, analytics, data-driven validation, and machine learning.

## Data Scientist / Data Analyst / Lead Developer

Talus Analytics (Boulder, CO)

01/18–04/19

Focused on integrating data analysis, visualization, and design.

Responsible for the full-stack development & technical direction of several analytics dashboards with emphasis on public health and global security.

## Data Scientist / Software Engineer

Nidhogg Consulting (Boulder, CO)

03/16–01/21

Co-founded as Chief Data Scientist and lead Software Engineer.

Responsibilities included project management, full stack web development, data science, and new contract acquisition.

## Applied Math Research Assistant

University of Colorado, Boulder (Boulder, CO)

09/15–05/17

Used machine learning models on gathered experimental lab data to predict data quality.

Developed and implemented a massively parallel simulation framework to analyze the behavior of soliton gasses in one-dimension systems.

Presented at the 2016 & 2017 Society for Industrial and Applied Mathematics (SIAM) student conferences

## Data Management Assistant


National Snow and Ice Data Center (Boulder, CO)

08/12–05/13

Large scale ETL tasks involving terabytes of research data and heavy applications of optimized Python data ingestion.

## Contact

 Denver, Colorado

 (719) 233 3514

 zoe@dataleek.io

 github/thedataleek

## </> Programming

- Python
- SQL
- Javascript / HTML / CSS
- Regular Expressions
- Bash / Shell Scripting

## ⚙️ Tools

- Machine Learning, Classification, Regression, Anomaly Detection
- Natural Language Processing (NLP), Large-Language Models (LLMs), Retrieval-Augmented Generation (RAG), Model Context Protocol (MCP), Agentic AI
- Amazon Web Services (AWS), Google Cloud Platform (GCP), Databricks, Prefect, Snowflake
- Polars, Scikit-Learn, Keras, PyTorch, Tensorflow, Pandas, Numpy, Scipy, NLTK, Flask, Asyncio, Jupyter, Git
- Matplotlib, Holoviews, D3, Graphviz, Looker, PowerBI
- CI/CD, Agile Development, Unix, Docker

## 🎓 Education

Applied Math B.S  
Computer Science Minor  
University of Colorado  
Boulder  
May, 2018

# ZOE FARMER

ML, Modeling, & AI Engineer

## Summer Internships

### Data Scientist

Khan Academy

2016

Built a probabilistic and Markov-chain network infection model for the interview ([github.com/thedataleek/kahn-interview](https://github.com/thedataleek/kahn-interview)). Developed new features, enhanced metrics, and maintained [khanacademy.org](https://khanacademy.org). Used investigative analytics to research SAT test scores. Developed a generalized learning model to assist with student growth and projection estimates for the SAT preparation tool.

### Data Scientist

Shutterstock

2015

Created an Apache Spark application to handle large scale ETL tasks on top of Shutterstock's Hadoop platform, as well as deployed and developed in-house big-data solutions

### Data Scientist

SpotXChange

2014

Developed an automatic Tableau Server reporting tool based on the semi-public API. Analyzed time-series data in an effort to identify similarities and overlapping data. Created a probabilistic model to analyze potential single-threaded scheduler optimizations

### Data Visualization

MITRE

2013

Created network data visualizations based on packet transfer data using common visualization libraries and tools. Responsible for challenges and solution implementations for the 2013 security Capture the Flag event.

## Projects

- PyColorado 2019 Board
- Boulder Python 2016-2019
- Black Arrow CLI Tool
- Discord Chat Plugins
- Hackathon Winner SQL/GPT4
- D3 Framework

## Hobbies

Magic the Gathering  
Diving deep into fantasy novels  
Gaming  
Pottery  
Drawing  
Puzzles



References Available Upon Request