

# ZOE FARMER

Data Scientist / Machine Learning Engineer

**Data Scientist / Machine Learning Engineer** 04/21–Current  
Square (San Francisco, CA)

*Built, tuned, and deployed machine learning models to support high-value stakeholder initiatives such as product value and anomaly detection.*

*Established a mature data ecosystem to support modeling work.*

*Supported a large product team with a wide array of data science and modeling ad-hoc requests.*

*Managed contracted roles to provide additional team bandwidth.*

**Data Scientist** 04/19–04/21  
Red Dot Storage (Louisville, CO)

*Built, tuned, and deployed key financial models to optimize the core revenue management platform.*

*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** 01/18–04/19  
Talus Analytics (Boulder, CO)

*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** 03/16–01/21  
Nidhogg Consulting (Boulder, CO)

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

*Responsibilities included project management, full stack web development, data science projects, and new job acquisition.*

**Applied Math Research Assistant** 09/15–05/17  
University of Colorado, Boulder (Boulder, CO)

*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** 08/12–05/13  
National Snow and Ice Data Center (Boulder, CO)


*Large scale ETL tasks involving terabytes of research data.*

*Involved heavy applications of optimized Python data ingestion along with bash scripting.*

## 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)

Amazon Web Services (AWS), Google Cloud Platform (GCP), Databricks, Prefect, Snowflake

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

## Data Scientist Internship

Khan Academy

*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.*

2016

## Projects

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

## Data Scientist Internship

Shutterstock

*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*

2015

## Hobbies

Sewing  
Dungeons & Dragons  
Martial Arts  
Video Games

## Data Scientist Internship

SpotXChange

*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*

2014

## Data Visualization Internship

MITRE

*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.*

2013

References Available Upon Request