ZOE FARMER

Data Scientist / Machine Learning Engineer

Data Scientist / Machine Learning Engineer
04/21-
Square (San Francisco, CA)

- *Built, tuned, and deployed to production an anomaly-detection model for vast amounts of device telemetry data across hundreds of KPIs to shorten critical-issue response time. This model was then integrated into Slack for DRI-alerting & metric visualization via custom-built reports with refined data visualizations to assist in triaging and decision-making abilities for product owners.*

- *Participated in and won "Developer's Choice" for a hack week in December 2023 for building a tool that could leverage GPT-4 to synthesize complex SQL answers via integrations of table metadata along with on-demand internal documentation.*

- *Established a mature data pipeline to support modeling work integrating between databases, data dictionaries, and user-generated content.*

- *Build, tuned, and deployed device & merchant cohort lifetime-value (LTV) models to support high level product owner decisions.*

- *Supported a large product team with a wide array of data science and modeling ad-hoc requests while building a comprehensive self-serve knowledge store through strict internal documentation processes while managing contracted roles to assist in demand.*

- *Independently gathered requirements, built modern data pipelines, and tackled high-priority issues regarding data quality & consistency.*

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

- *Built, tuned, and deployed two key financial models to optimize the core revenue management platform via precise current-tenant gradual rate increases and realtime market-value rates for new-tenants.*

- *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.*

- *Leveraged modern market analysis, business research, geospatial modeling, regression analysis, and performance optimizations*

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 of several analytics dashboards with emphasis on public health and global security.*

- *Drove technical focus and direction surrounding data & web-development platforms using modern libraries and philosophies.*

- *Pioneered a key-rewrite of the system architecture to fully leverage ES6 Promises to allow for a more responsive user experience.*

- *Created and presented at local meetups on an extensible open-sourced D3 framework allowing for fast development of intricate visuals.*

Contact

- Denver, Colorado
- (719) 233 3514
- zoe@dataleek.io
- github/thedataleek

Programming

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

Tools

Mathematics, Probability, Statistics

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, Fivetran

Pandas, Pola.rs, Numpy, Scipy, NLTK, Flask, Multiprocessing, Multithreading, Asyncio, Jinja Templating, Jupyter, Git

Matplotlib, Seaborn, Holoviews, D3, Graphviz, Looker, PowerBI

Software Engineering, CI/CD, Agile Development, Linux, Unix, MacOS, Docker

Education

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

ZOE FARMER

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.

⬡ Projects

- PyColorado 2019 Board
- Boulder Python 2016-2019
- Black Arrow CLI Tool
- Discord Chat Plugins

⬡ Hobbies

Sewing
Dungeons & Dragons
Martial Arts
Video Games

Summer Internships

Data Scientist Intern
2016

Khan Academy

Built a probabilistic and Markov-chain network infection model for the interview (github.com/ thedataleek/kahn-interview). Developed new features, enhanced metrics, and maintained 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 Intern
2014

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

Data Scientist Intern
2015

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

Data Visualization Intern
2013

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.