

# Matthew Xie

mkxie@berkeley.edu | (925) 997-1661 | github.com/mkxie | linkedin.com/in/matthewkxie/

## EDUCATION

**University of California, Berkeley**

**August 2017—May 2021**

**BA, Computer Science** | **GPA: 3.73**

**Coursework:** Data Structures, Designing Information Systems and Devices, Discrete Mathematics and Probability Theory, Principles and Techniques of Data Science, Machine Structures, Efficient Algorithms and Intractable Programs, Introduction to Database Systems, Operating Systems and Systems Programming

**Clubs and Activities:** Pioneers in Engineering (PiE), Berkeley Engineers and Mentors (BEAM), AFX Dance

## EXPERIENCE

**Data Science Intern**—*Sandia National Laboratories, Livermore, CA*

**May 2019—August 2019**

- Utilized Python and Apache Spark on Sandia's High Performance Data Analytics clusters to scale up distributed data processing on machine learning tasks and demonstrate big data applications for new users
- Built a convolutional neural network classifier for deep learning on tens of thousands of seismic records to distinguish earthquakes from man-made quarry blasts and gain actionable insights for future inferences. Achieved 97% accuracy using Spark and Keras leveraging multiple GPUs for training and parallelizing evaluations for new inferences
- Wrote tutorials and guides to introduce new users to machine learning workflows using Spark and Keras on Sandia's data analytics clusters, including setting up Spark, configuring virtual environments for both CPU and GPU architectures, distributing work, and monitoring job status

**Software Developer**—*Pioneers in Engineering, Berkeley, CA*

**September 2018—May 2019**

- Implemented an internal staff check-in database to keep track of attendance during worksessions using Python, SQLite, and Google Cloud's NoSQL-based Datastore
- Maintained and expanded features for student worksessions web app built with Ruby on Rails using Google Sheets API, Slack API, Flask, and Python to check staff availability and automate sending daily Slack notifications to assigned staff

**Engineering Intern**—*Lawrence Livermore National Laboratory, Livermore, CA*

**June 2018—August 2018**

- Automated installation and configuration processes on Linux-based HPC clusters using Ansible (cmd/Bash)
- Worked with a team of four to add new functionality to a Python software suite based off of the open-source Pyvmomi library for virtual machine automation and orchestration. Improved efficiency by eliminating the need to do tasks through VMware's ESXi GUI, and presented to industry experts for feedback (Python)

**Academic Intern**—*UC Berkeley Division of Data Sciences, Berkeley, CA*

**January 2019—May 2019**

- Provided TA assistance during lab sections for UC Berkeley's Foundations of Data Science course (DS 100)
- Helped students understand challenging concepts of data science during labs, homeworks, and projects

## PROJECTS

**Concurrent Cached Filesystem** (Go)

**April—May 2019**

- Built a cached file system to handle HTTP requests and return appropriate files or errors
- Implemented Go concurrency for multithreaded execution of file reads from disk and placed them in the cache for faster file handling

**Bay Area Maps** (Java)

**April 2018**

- Provided the back end for a Google Maps-like application for the East Bay Area, including zoom capabilities, rendering map images from user queries, and finding shortest routes using A\* search algorithm

## SKILLS

**Languages:** Java, Python, C, Golang, HTML/CSS, Ruby/Rails, SQL

**Technologies:** Git, NumPy, Pandas, Spark, TensorFlow, Keras, Matplotlib, Seaborn, Flask, SQLite, Docker, Linux, HPC