MATTHEW XIE

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

EDUCATION

University of California, Berkeley **August 2017—December 2020**
BA Computer Science **|** ***GPA***: 3.73
Coursework: Data Structures, Algorithms, Discrete Math and Probability Theory, Data Science, Machine Structures, Designing Information Systems and Devices, Operating Systems*, Databases*, Software Engineering**, Security**
*In Progress **Planned for Spring 2020
Clubs and Activities: Pioneers in Engineering, Berkeley Engineers and Mentors, AFX Dance

EXPERIENCE

Data Engineering R&D 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 with Spark and Keras for deep learning on tens of thousands of seismic records to distinguish earthquakes from man-made quarry blasts with 97% accuracy and gain actionable insights for future inferences

Data 100 Academic Intern—*UC Berkeley Division of Data Sciences, Berkeley, CA* **January 2019—May 2019**
- Assisted a 30 person section on challenging concepts of data science such as Pandas, SQL, data visualization, regression, feature engineering, and classification during labs, homeworks, and projects

Software Developer—*Pioneers in Engineering, Berkeley, CA* **September 2018—May 2019**
- Implemented an internal staff check-in database to keep track of attendance for 200+ members during worksessions using Python, SQLite, and Google Cloud's NoSQL-based Datastore
- Maintained and expanded features for student worksession web app serving 20+ high schools built with Ruby on Rails using Google Sheets API, Slack API, Flask, and Python to check staff availability and automate daily Slack notifications

Engineering Intern—*Lawrence Livermore National Laboratory, Livermore, CA* **June 2018—August 2018**
- Automated installation and configuration processes on Linux-based HPC clusters using Ansible
- Collaborated with a team of four to add new functionality to an open-source Python software suite for virtual machine orchestration for LLNL's Livermore Computing division and Systems Administration Group
- Improved efficiency by eliminating the need to control VMs through VMware's ESXi GUI, and presented to industry experts for feedback

PROJECTS

Bloggit (Flask, SQLite)—*github.com/mkxie/bloggit* **April 2019—Present**
- Constructed a full-featured Reddit-like blog web application where users can create, update, and destroy accounts and blog posts and interact with each other on the website

Concurrent Cached Filesystem (Go) **April 2019—May 2019**
- Implemented Go concurrency on a cached file system for multithreaded execution of file reads from disk and placing files in the cache for faster handling

Spam Email Classification (Python) **October 2018—November 2018**
- Created a logistic regression model using Pandas and Scikit-learn with NLP techniques to classify spam vs ham emails with 90% accuracy

Bay Area Maps (Java) **April 2018**
- Provided the backend 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: Python, Java, C, Golang, SQL, HTML/CSS, Ruby/Rails
Technologies: Git, Pandas, Spark, Keras, TensorFlow, Flask, SQLite, Docker, Linux, HPC