# Jacob Cutter

Mountain View, CA | 510-861-0706 | jecutter92@gmail.com
linkedin.com/in/jacob-cutter/ | **jecutter.github.io/** | **github.com/jecutter**

## SKILLS

**Programming Languages:** Python, C++, BASH/shell, SQL
**Tools:** Git, Scikit-Learn, PyTorch, Keras, Pandas, NumPy, SciPy, Matplotlib, Seaborn, Bokeh, Flask
**Techniques:** Machine learning (supervised, unsupervised, deep learning), NLP (BoW, TF-IDF, sentiment analysis), signal processing, statistical modeling and analysis, end-to-end data pipelining, parallel computing

## EXPERIENCE

**Data Scientist,** *Deepgram,* Mountain View CA — Oct 2020 - Present

- Maintaining and updating state-of-the-art deep learning ASR models
- Reporting KPIs for core ASR models and developing metrics to track customer success
- Developing data analysis and model training tools that use parallel and multi-GPU processing
- Procuring, cleaning, and pre-processing speech data for model training and evaluation
- Curating terabytes of audio data, including automating Q.A. and classification processes
- Syncing with Data Ops, Research, and Engineering teams to coordinate cross-functional efforts

**Data Science Fellow,** *Insight,* San Francisco CA — May 2020 - July 2020

- Created a music classification app for listeners and content creators to filter songs by emotion
- Combined emotional labels mined from Last.fm SQLite databases, Spotify audio features, and song lyric sentiment to build classifiers that predict the emotional charge of songs with up to 71% accuracy
- Deployed the classification models on AWS in an interactive Streamlit web application

**Graduate Student Researcher and T.A.,** *Physics Department,* UC Davis — Sept 2014 - August 2020

- Spearheaded design and operation of local R&D particle detection experiments to characterize important nuclear processes and improve statistical models used in dark matter searches
- Developed custom C++ and Python software for signal processing, data reduction and visualization, synthesizing TBs of noisy waveform data into physical measurements
- Used MySQL database replication and Flask to remotely monitor lab operations via web interface
- Managed large-scale data processing pipelines for an international multi-million dollar experiment
- Taught and devised lesson plans for discussion/lab sections for multiple undergraduate physics courses, receiving excellent student evaluations averaging higher than 4.5/5 across the board
- Mentored ~75 high school students over 3 summers as part of the COSMOS program, working with students to execute intensive data-driven astrophysics projects in just 1 month

## PROJECTS

**NBA Basketball Analytics |** github.com/jecutter/nba-data-models, jecutter.github.io/blog/ — Jan 2020

- Web-scraped several seasons of NBA player, lineup, and play-by-play data using Scrapy and Selenium
- Created an interactive visualization dashboard using Bokeh to explore player and lineup data
- Used a Random Forest model to perform useful stylistic player classifications with 92% accuracy
- Developed an RAPM model using ridge regression on matchup results harvested from play-by-play data, providing a lineup-independent impact metric for scouting undervalued players

## EDUCATION

**University of California, Davis** — August 2020
Ph.D. in Experimental Particle Physics, Designated Emphasis in Nuclear Science

**University of California, Davis** — June 2014
B.S. in Particle Physics