

# Jacob Cutter

San Francisco, CA | 510-861-0706 | [jecutter92@gmail.com](mailto:jecutter92@gmail.com)

[linkedin.com/in/jacob-cutter/](https://www.linkedin.com/in/jacob-cutter/) | [jecutter.github.io/](https://github.com/jecutter) | [github.com/jecutter](https://github.com/jecutter)

## SKILLS

---

**Programming Languages:** Python, C++, BASH/shell, SQL

**Tools:** Git, Scikit-Learn, Pandas, NumPy, SciPy, Matplotlib, Seaborn, Bokeh, Flask, Streamlit

**Techniques:** Machine learning (regression, classification, clustering), NLP (BoW, TF-IDF, sentiment analysis), signal processing, statistical modeling and analysis, end-to-end data pipelining, parallel computing

## EXPERIENCE

---

**Data Science Fellow, *Insight*, San Francisco CA**

May 2020 - Present

- Created a music classification app for listeners and content creators to filter songs by emotion
- Combined Spotify audio features, sentiment analysis of song lyrics, and emotional labels extracted from Last.fm SQLite databases to build Logistic Regression and Random Forest classification models that predict the emotional charge of songs with up to 71% accuracy
- Deployed the model on AWS as an interactive Streamlit web application (hosted at [datapoint.xyz](https://datapoint.xyz))

**Graduate Student Researcher, *Physics Department*, UC Davis**

Sept 2014 - July 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, resulting in the synthesis of 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 pipelines as Data Processing Manager for a multi-million dollar experiment and gave verbal bi-weekly reports to an international collaboration of scientists
- Created an algorithm for 3D position reconstruction of particle interactions, combining multiple spatial interpolation techniques to minimize systematic errors and maximize potential for particle discovery

**Teaching Assistant, *Physics Department*, UC Davis**

Fall 2014 - Summer 2016

- Devised lesson plans and taught 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](https://github.com/jecutter/nba-data-models), [jecutter.github.io/blog/](https://jecutter.github.io/blog/)

Jan 2020

- Used Scrapy and Selenium to scrape several seasons of player, lineup, and play-by-play data
- 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