

# Min Gu (Min) Jo

(510) 365-4988 | [mingu08@berkeley.edu](mailto:mingu08@berkeley.edu)  
<http://mingujo.com> | GitHub : [github.com/mingujo](https://github.com/mingujo)

## EDUCATION

### University of California, Berkeley

- **B.A. in Computer Science, Statistics, and Economics**

- Cumulative GPA : 3.42/4.00, Expected Graduation : December 2016

*Relevant Coursework:* Machine Learning • Software Engineering • Statistical Inference & Computing • Database • Linear Modeling • Artificial Intelligence • Computer Architecture • Internet Networks\*

## SKILLS

### Programming Languages

- Python, Java, R, SQL, C

### Framework, Library, Database, Platform

- Flask, Django, TensorFlow, Scikit-Learn, Keras, Heroku, Celery, Boto (AWS EC2), Postgres

## PROFESSIONAL EXPERIENCE

June 2016 -  
August 2016  
Berkeley, CA

### LeadGenius

#### Software Engineer Intern

- Deployed a Flask web app which automates labeling of B2B sales outreach email responses using a Recurrent Neural Network model with LSTM architecture built with Deep Learning lib, TensorFlow
- Developed a system which automatically notifies clients of new positive outreach responses
- Implemented an automatic migration of all labeled email threads from Nylas to Postgres database
- Set up a Celery task scheduler which automatically launches and destroys a GPU-enabled Amazon EC2 instance using Boto to retrain a new learning model weekly
- Predicted the open rate of outreach emails based on subject-line contents using multivariate linear regression

January 2016 -  
Present  
Berkeley, CA

### Berkeley Institute of Data Science | OskiLab

#### Undergraduate Researcher

- Work with a Ph.D. Student to explore a number of phenomena, including how consumers and producers understand and evaluate products in murky state-legal marijuana markets
- Use Keras (deep learning library for Theano) to detect article text bodies in Bon Appetit magazines
- Scraped user reviews on *weedmap.com* using web browser automation tool, Selenium

## RESEARCH/PROJECT EXPERIENCE

Spring 2016

### Kaggle Challenge: Rossmann Drugstore Store Sales Prediction

- Used 3 different machine learning algorithms to forecast drugstore daily sales : multivariate linear regression, Random Forest regression, and Gradient Boosting with regression trees

Fall 2015

### Detection of Activated Brain Regions Under Mixed Gamble Task (In-class Project)

- Investigated the relationship between the brain activity and the behavior of the subjects towards the 50/50 gambling situations using a whole-brain robust regression analysis
- Preprocessed and analyzed fMRI image data to identify active regions of the participants' brains

Fall 2015

### Kaggle Challenge: Bag of Words Meets Bags of Popcorn

- Applied text analysis(NLP) methods of TFIDF vectorizer and Google's word2vec on IMDb movie reviews to perform sentiment analysis (96% accuracy | top 11th percentile when submitted)

Fall 2015

### Kaggle Challenge: Titanic Disaster Survival Prediction

- Developed machine learning ensemble models to predict the survival in Titanic disaster (80% accuracy when submitted)

Spring 2013

### Probabilistic Modeling of Interactions on UC Berkeley Campus

#### Prof. David Aldous: Undergraduate Research Group

- Designed an independent research topic to predict and visualize common routes of UC Berkeley undergraduates with different majors and their interactions on campus
- Collected survey data from 130+ undergraduates across 5 different majors on MySQL database
- Presented to 20+ undergraduates at the Statistics Undergraduate Research Poster Session

Spring 2013

### Prediction of Kobe Bryant's Performance in His Next Game (In-class Project)

- Scraped Kobe's seasonal data from *basketball-reference.com* and selected relevant predictors
- Applied regression analysis and feature shrinkage methods to create statistical models for prediction and cross validated to evaluate the different machine learning models using R