Molly Can Zhang

217-417-7857
mollyzhang@ucsc.edu

Education

2016–present **PhD(expected), Computer Science**, *University of California, Santa Cruz*, Santa Cruz, *CA*.
2014–2016 **MS, Computer Science**, *University of California, Santa Cruz*, Santa Cruz, *CA*.
2009–2011 **MS, Biochemistry**, *University of Illinois at Urbana-Champaign*, Urbana, *IL*.
2005–2009 **BS, Bioengineering**, *Beijing Institute of Technology*, Beijing, *China*.

Skills

Languages **Python**, **C**, Shell Script, Matlab, Java, JavaScript
AI **Machine Learning**, **Deep Learning**, Convolutional Net, LSTM, Word2Vec, Autoencoder
Genomics **RNA-seq analysis**, Sequence Alignment, Samtools, GATK
Web HTML/CSS, **Web2py**, **Bootstrap**, React, jQuery, Reactive.js
Others **Git**, **APIs**, Unix/Linux, Algorithms

Experience

2018 summer **Applied Scientist Intern**, *Amazon*, Boston, MA.
Improved Alexa Natural Language Understanding pipeline by double-digit percentage in three months by implementing latest research papers and models, such as word2vec word embedding, phrase representation and phrase embedding, weighted utterance embedding and deep recurrent neural network (LSTM).

2017 summer **Data Science Intern**, *Asana*, San Francisco, CA.
Developed and productionized unsupervised machine learning pipeline to cluster Asana premium users into meaningful sectors with distinct behaviors. Created visualization tool to allow for selecting and highlighting users and user clusters by company, industry, job title and location.

2014-Present **Graduate Student Researcher**, UC Santa Cruz, CA.
Developed multi-task deep learning methods to use multi-omics data to predict cancer patients' survival outcome and extracted accurate low-dimension representation of high-dimensional genomics data with autoencoders. Created the world largest public database containing pathogenic BRCA variants and hosted data at `brcaexchange.org` for free.

2012-2014 **Research Associate**, *Genentech, Inc*, South San Francisco, CA.
Formulation development and expiration dating of monoclonal antibody cancer drugs.

2011-2012 **Co-founder**, *VirtualFit*, Urbana, IL.
Cofounded online clothing shopping startup to build 3D models of human body from 2d Images and to visualize person wearing clothing of interest.

Projects

October 2016 **Natural language processing of US presidential election**.
Auto-generates president candidate's debate speech with LSTM recurrent neural network model

June 2016 **BetaGo - Predicting professional moves in Go games**.
Trained convolutional neural network to predict Go moves from professional players from 17 million training examples parsed from 85,000 professional Go game record, achieved comparable accuracies to AlphaGo

December 2015 **Flurbo - a personal budgeting web application**.
Turn personal finance spreadsheet to web application with web2py, Ractive.js, d3 and Bootstrap. Flubo can automatically generates weekly budgets and monthly financial report along with spending visualization

Other

Marathon Finisher, Classical Guitar Player, Internship Applicant