

ZHENGHAO CHEN

MACHINE LEARNING ENGINEER

zhenghao@z-chen.net | www.z-chen.net | SF Bay Area

EDUCATION

M.S. Computer Science

(AI and Biocomputation)

GPA: 4.13

Stanford University 2013

B.S. Computer Science

(with distinction, honors & minor in Math)

GPA: 4.04

Stanford University 2013

Awards

Frederick E Terman Engineering Award
(top 5% of Stanford Engineering seniors)

President's Award for Academic Excellence in the Freshman Year (top 3% in class)

Relevant Coursework

Machine Learning, Statistical Learning Theory, Probabilistic Graphical Models, Theory of Statistics, Convex Optimization, Information Theory

SKILLS

Languages

Native proficiency in Mandarin Chinese and English (speaking, reading, writing)

Technical Skills

Proficient in Python, Javascript, Scala. Some experience working in C++ and Java.

Familiar with modern ML frameworks (Tensorflow, JAX). Experienced in applied ML research, data-infrastructure engineering and building data products.

EXPERIENCE

Principal ML Engineer

Calico Labs | 2017 - Now

Built state-of-the-art models for modeling cell state in small molecule and genetic screens. Deployed ML models for prioritization and selection of perturbations for large-scale screening programs.

Developed novel methods and models to:

- Extract age-related features from high-dimensional time series data from longitudinal studies.
- Analyze in-situ multiplexed cell imaging for studying the cellular microenvironment in various contexts such as cancer tumor microenvironments or tuberculosis.
- Achieve state-of-the-art performance on general MHC binding prediction.

Senior Data Scientist

Coursera | 2012 - 2017

Founding member of the analytics team and helped grow team to 20 data scientists / engineers

Built key data infrastructure such as the in-house A/B testing platform which supported hundreds of experiments annually and the content recommendation system powering all aspects of content discovery on Coursera.

Led the development of data products such as the onboarding content discovery experience, weekly recommendation email, recommendations for course sequences.

Initiated and led research efforts to use machine learning to improve learning experience. Developed ML models for debiasing peer grading, automated feedback for programming assignments, recommendation models for courses / course sequences etc.

Designed, carried out and published large scale studies on the effectiveness of online instruction and instructional design as well as the long-term benefits of MOOCs on learners.

Intern

Counsyl | June - Sept 2012

Improved error correction and accuracy of in-house Fragile X screening model.

Research Assistant

Stanford AI Labs | 2009 - 2012

Applied probabilistic modelling methods to computational cancer biology and biomedical imaging with Prof. Daphne Koller.

Developed a method for automatic feature extraction of image features from whole slide images of brain tumors. Identified new subtypes of glioblastoma multiforme with different survival and treatment characteristics.

Developed models for unsupervised feature learning using deep neural networks for image recognition and speech recognition with Prof. Andrew Ng. Published several papers in ICML, NeurIPS.

PUBLICATIONS AND TALKS

- **Learning good representations of cell state from cell painting**, [Z.Chen](#), C Jan, F Li, J Xu. NeurIPS Learning Meaningful Representations of Life Workshop 2019
- **Modeling multiplexed images with Spatial-LDA reveals novel tissue microenvironment**, [Z.Chen](#), I Soifer, H Hilton, L Keren, V Jojic. I CML Comp Bio Workshop 2019
- **Composition and decomposition of GANs**, YC Harn, [Z.Chen](#), V Jojic. arXiv Jan 2019.
- **Real-time programming exercise feedback in MOOCs**, [Z.Chen](#), A Nguyen, A Schlender, J Ngiam. EDM 2017
- **Mapping the Pairwise Choices Leading from Pluripotency to Human Bone, Heart, and Other Mesoderm Cell Types**, K Loh, A Chen, PW Koh, T Deng, R Sinha, J Tsai, A Barkal, K Shen, R Jain, R Morganti, SC Ng, N Fernhoff, B George, G Wernig, R Salomon, [Z.Chen](#), H Vogel, J Epstein, A Kundaje, W Talbot, P Beachy, LT Ang, I Weissman. Cell July 2016.
- **Who's Benefiting from MOOCs and Why?**, [Z.Chen](#), B Alcorn, G Christensen, N Eriksson, D Koller, EJ Emanuel. Harvard Business Review September 2015.
- **What Matters, What Doesn't? A Coursera-Wide Look at Course Metrics**, N Eriksson, A Parisi-Amon, [Z.Chen](#). Talk at 2015 Coursera Partners Conference, Irvine, CA.
- **Dissecting an Online Intervention for Cancer Survivors**, [Z.Chen](#), PW Koh, P Ritter, K Lorig, E Bantum, S Saria. Health Educ. Behavior. 2014.
- **Retention and Intention in Massive Open Online Courses**, D Koller, A Ng, CB Do, [Z.Chen](#). Educause Review June 2013.
- **Self-Driven Mastery in Massive Open Online Courses**, CB Do, [Z.Chen](#), R Brandman, D Koller. MOOCs Forum September 2013.
- **Automated Population-scale Screening for Fragile X: Validation and Experience on 76,421 Samples**, SA Patterson, MR Theilmann, [Z.Chen](#), IS Haque. Poster at AMP 2013
- **Tuned Models of Peer Assessments in MOOCs**, C Piech, J Huang, [Z.Chen](#), CB Do, A Ng, D Koller. EDM 2013
- **A Graph Regularization Based Approach for Gene Scoring**, [Z.Chen](#). Stanford CS Honors Thesis 2012.
- **Sparse Filtering**, J Ngiam, PW Koh, [Z.Chen](#), S Bhaskar, A Ng. NIPS 2011. Spotlight paper.
- **Learning Deep Energy Models**, J Ngiam, [Z.Chen](#), PW Koh, A Ng. ICML2011.
- **On Random Weights and Unsupervised Feature Learning**, A Saxe, PW Koh, [Z.Chen](#), M Bhand, B Suresh, A Ng. ICML 2011.
- **Tiled Convolutional Neural Networks**, QV Le, J Ngiam, [Z.Chen](#), D Chia, PW Koh, A Ng. NIPS 2010.
- **Lower Bound on the Time Complexity of Local Adiabatic Evolution**, [Z.Chen](#), PW Koh, Z Yan. Phys. Rev. A 74, 2006.