

# CALVIN CHI

---

CONTACT	calvin.t.chi@gmail.com (626) 203-1829	<a href="https://github.com/CalvinTChi">https://github.com/CalvinTChi</a> <a href="https://calvintchi.github.io">https://calvintchi.github.io</a>
EDUCATION	<b>University of California, Berkeley</b> Ph.D., Computational Biology	2015 - 2020
	<b>Case Western Reserve University</b> B.S., Biochemistry <i>summa cum laude</i> (GPA 4.0/4.0)	2011 - 2014
INDUSTRY	<b>Applied Scientist</b>	Aug 2020 - present
	<b>Applied Scientist Intern at Amazon</b> • Developed LSTM-based deep learning model for credit abuse classification for Amazon Business, estimated to reduce credit write-off loss by 20%. • Shipped ~ 1,000 lines of python and SQL code for data and model production.	May 2019 - Aug 2019
PUBLICATIONS	Chi, Calvin, et al. Bipartite graph-based approach for clustering of cell lines by gene expression-drug response associations. <i>Bioinformatics</i> . [In Review]	
	Chi, Calvin. HLA Allele Imputation with Deep Convolutional Neural Network. <i>Bioinformatics</i> . [In Preparation]	
	Chi, Calvin, et al. Epigenetic stratification identifies clinically relevant disease subgroups in Sjögren's syndrome with differential genetic risk at the major histocompatibility complex. <i>Nature Medicine</i> . [In Preparation]	
	Chi, Calvin, et al. Hypomethylation of Immune Genes Mediates Methylation Quantitative Trait Loci at the Major Histocompatibility Complex in Sjögren's Syndrome. <i>PLoS genetics</i> . [In Preparation]	
	Chi, Calvin, et al. "Admixture mapping reveals evidence of differential multiple sclerosis risk by genetic ancestry." <i>PLoS genetics</i> 15.1 (2019): e1007808.	
PROJECTS	<b>Embedding-Augmented Deep CNN for PubMed Journal Recommendation</b> • Journal detection from PubMed abstract with 415,381 programmatically-collected abstracts. • Compared multitask and embedding-augmented CNNs with output space of 1,548 journals. • Best performance when CNN input augmented with topic and impact factor embeddings, with accuracy 23.7% and 90% of true journals in top 60 recommendations.	Dec 2018
	<b>Data Augmentation using GAN for Breast Cancer Classification</b> • Synthetic data augmentation using DCGAN to improve histology breast cancer classification with Resnet-18 re-trained on 5,547 breast histology images. • Augmentation with 400 DCGAN images improved prediction accuracy and precision by 5% and 12% respectively, but decreased recall by 15%.	May 2018
SKILLS	<b>Programming:</b> Python R, SQL, Java, Bash, Matlab, HTML, CSS, Javascript, C <b>Libraries:</b> Scikit-Learn, H2O, Keras, TensorFlow, PyTorch	
AWARDS	NSF Graduate Research Fellowship	Mar 2017