CALVIN CHI

CONTACT	calvin.t.chi@gmail.com	https://github.com/CalvinTChi
	(626) 203-1829	https://calvintchi.github.io

EDUCATION

University of California, Berkeley 2015 - 2020
Ph.D., Computational Biology

Case Western Reserve University 2011 - 2014
B.S., Biochemistry
summa cum laude (GPA 4.0/4.0)

INDUSTRY

Applied Scientist Aug 2020 - present

Applied Scientist Intern at Amazon May 2019 - Aug 2019
- Developed LSTM-based deep learning model for credit abuse classification for Amazon Business, estimated to reduce credit write-off loss by 20%.
- Shipped $\sim 1,000$ lines of python and SQL code for data and model production.

PUBLICATIONS

Chi, Calvin, et al. Bipartite graph-based approach for clustering of cell lines by gene expression-drug response associations. *Bioinformatics.* [*In Review*]

Chi, Calvin. HLA Allele Imputation with Deep Convolutional Neural Network. *Bioinformatics.* [*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. *Nature Medicine.* [*In Preparation*]

Chi, Calvin, et al. Hypomethylation of Immune Genes Mediates Methylation Quantitative Trait Loci at the Major Histocompatibility Complex in Sjögren's Syndrome. *PLoS genetics.* [*In Preparation*]

Chi, Calvin, et al. "Admixture mapping reveals evidence of differential multiple sclerosis risk by genetic ancestry." *PLoS genetics* 15.1 (2019): e1007808.

PROJECTS

Embedding-Augmented Deep CNN for PubMed Journal Recommendation Dec 2018
- 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.

Data Augmentation using GAN for Breast Cancer Classification May 2018
- 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%.

SKILLS

Programming: Python, R, SQL, Java, Bash, Matlab, HTML, CSS, Javascript, C
Libraries: TensorFlow, PyTroch, Keras, Scikit-Learn, H2O

AWARDS

NSF Graduate Research Fellowship Mar 2017