XIN CHENG

EMAIL: xin0558@gmail.com **PHONE:** 706-201-9416 **LOCATION:** San Jose, CA 94088
LINKEDIN: xincheng0558

SKILLS

PROGRAMMING LANGUAGE: Python, sql
PACKAGES: sklearn, matplotlib
SUPERVISED ML: Linear regression, Tree based methods, Ensemble methods
UNSUPERVISED ML: Natural language processing, Clustering
DATA/CLOUD: AWS, sql, MongoDB
VISUALIZATION TOOLS: Tableau

EXPERIENCE

Metis, *Data Scientist*, San Francisco, CA — Mar. 2020 - June 2020

- Metis is an ACCET accredited 12 week immersive data science bootcamp focused on project oriented learning
- The core curriculum is centered around Python, statistics, supervised and unsupervised machine learning, exploratory data analysis, databases, and visualization techniques
- Completed five self designed data science projects from conception to presentation; including data collection, data management, exploratory data analysis, modeling, and visualizations

Synthego, *Production Scientist*, Redwood City, CA — Apr. 2019 - Current

- Analyzed cell line optimization data with pandas
- Finished over 10 advanced gene editing projects using CRISPR with automation

Senior Research Associate, Redwood City, CA — Nov. 2017 - Apr. 2019

- Designed and validated Hamilton liquid handling automation methods for production
- Helped build production team from ground up

Applied Stemcell, *Research Associate*, Milpitas, CA — July 2015 - July 2017

- Generated over 40 homozygous point mutation/deletion mammalian cell lines with CRISPR/Cas9.
- Overseeing teratoma projects as a project manager
- Transfected mouse embryo stem cells with electroporation
- Cultured hiPS cells

PROJECTS

Natural Language Processing on COVID-19 Publications

- Analyzed 10000 COVID-19 publications with Natural language processing, extracted topics with topic modeling, visualized topics with tSNE. This project simplified researcher's search for relevant publications

Bank Marketing Classification

- Analyzed bank marketing data, in order to classify which customer will sign up for term deposit, I did classification on an imbalanced dataset, compared XGBoost, random forest and linear regression to classify the target variable

EDUCATION

University of Georgia — Aug. 2011 - Dec. 2014
M.S. Biological Engineering 2014

Hunan University — Sept. 2009 - Sept. 2010
B.S. Biotechnology 2010