

# Daniel Lin

## Data Scientist

✉ danman99551@gmail.com  
☎ 949-878-0385  
📍 San Francisco, CA 94109  
in daniel-lin91  
🔗 dl91

## EDUCATION

University of Pennsylvania  
Post-Baccalaureate Cell Biology  
2016

Juniata College  
B.S. Biology 2014

## SKILLS

### PROGRAMMING LANGUAGES

Python  
SQL

### MACHINE LEARNING ALGORITHMS

Linear Regression  
Logistic Regression  
K-nearest Neighbors  
Support Vector Machine  
Naive Bayes  
Random Forest  
Gradient Boosting  
Natural Language Processing  
Clustering  
Neural Networks

### DATABASES

PostgreSQL  
MongoDB

### VISUALIZATION

Matplotlib  
Seaborn  
Tableau  
Excel/PowerPoint

### LIBRARIES

Pandas  
Numpy  
NLTK  
BeautifulSoup  
TensorFlow  
Keras  
Scikit-Learn

## EXPERIENCE

### Metis Data Science Bootcamp

Data Scientist

San Francisco, CA  
Mar. '20 - June '20

- Metis is an ACCET accredited 12-week immersive data science bootcamp focuses on Python, statistics, supervised and unsupervised machine learning, data analysis, data visualizations, database management, and communication
- Designed five end-to-end projects utilizing data acquisition, cleaning and modeling to derive data-driven insights, project highlights can be seen in the projects section
- Presented results to technical and non-technical audiences

### Unity Biotechnology

Research Associate II

San Francisco, CA  
Aug. '17 - Mar. '20

- Designed and implemented research experiments on diabetic retinopathy, and age-related macular degeneration
- Performed medium-high throughput drug screening on senolytics and observed efficacy in human retinal microvascular endothelial cells, and retinal epithelial cells
- Assisted in completing IND package for new compounds. Including TE/ME assay development
- Advanced cell culture experience in primary human cell line, cancer cell lines
- Developed gene expression and protein quantification pipeline
- Strong data analysis skills through Excel and Prism
- Working knowledge with TECAN, including designing and writing scripts

### Children's Hospital of Philadelphia/ University of Pennsylvania

Research Technician II/ Lab Manager

Philadelphia, PA  
May '15 - July '17

- Conducted research on Friedreich's Ataxia
- Performed tissue culture on patient derived fibroblast, iPS derived cardiomyocytes, muscle cells, lymphocytes, mouse fibroblasts
- Investigated the effect of random shRNAs on cell line by screening a shRNA library
- Designed drug screening experiments for potential therapeutic compounds in cells
- Executed functional and physiological experiments using various instruments including: SeaHorse, Gene5, Luminex, calcium flux analyses
- Assessed gene expression and protein quantification analyses through DNA/RNA extraction, cDNA generation, PCR, qPCR, Western Blot, ELISA
- Model organism development using morpholinos in Zebra Fish

## DATA SCIENCE PROJECTS

### Dog Breed Recognition System (Neural Networks)

- Built a convolutional neural network using Tensorflow and Keras capable of identifying the dog breed from an image
- Designed sequential convolutional neural networks and trained on dog images
- Used transfer learning to train the model on a dataset of dog images
- Achieved a 80% accuracy on the test dataset

### Job Recommender System (Natural Language Processing)

- Used MongoDB to manage data storage
- Identified topics from job descriptions using LSA, LDA, and NMF
- Visualized topics using PyLDAvis
- Utilizing NLP and topic modeling to build a job recommendation system for users

### Predicting Twitter User Gender (Classification Modeling)

- Stored data into AWS SQL Tables using PostgreSQL
- Employed NLP to process text messages
- Utilized classification models such as Random Forest, SVM, Logistic Regression, Naive Bayes, and KNN to predict the gender of twitter users based on their profile, tweets and retweets

### Predicting NBA Player Salaries (Linear Regression)

- Scraped data from websites using BeautifulSoup
- Built supervised models that predict NBA player salaries based on prior year performance
- Utilized linear regression to make predictions and L1 and L2 regression to prevent overfitting

### Exploratory Data Analysis on New York City Subway Activity

- Utilizing pandas to clean data and perform analysis on New York MTA turnstile data
- Visualized data using Matplotlib and Seaborn
- Determined stations with highest traffic to deploy fundraising team for a fictitious organization, WomenTechWomenYes

## PUBLICATIONS

### Ferroptosis in Friedreich Ataxia ·

M. Grazia Cotticelli, Shujuan Xia, Daniel Lin, Taehee Lee, Leila Terrab, Peter Wipf, Donna M. Huryn and Robert B. Wilson

Journal of Pharmacology and Experimental Therapeutics April 1, 2019, 369 (1) 47-54; DOI: <https://doi.org/10.1124/jpet.118.252759>

### Identification of p38 MAPK as a Novel Therapeutic Target for Friedreich's Ataxia ·

Cotticelli, M.G., Xia, S., Kaur, A. et al.

Sci Rep 8, 5007 (2018). <https://doi.org/10.1038/s41598-018-23168-x>