

# JULIAN CHENG

## DATA ANALYTICS

### CONTACT

---

✉ juliancheng2k@gmail.com  
☎ 781-353-9159

🐦 chengsaw1  
🌐 /in/julian-cheng-950a901b3/  
🌐 juBmam

### SKILLS

---

**PROGRAMMING LANGUAGES:** Python, Ruby, HTML  
**LIBRARIES:** Pandas, Numpy, SKLearn, Tweepy, Yellowbrick, Tensorflow, Keras  
**DATA VISUALIZATION:** Matplotlib, Seaborn, Flask, D3  
**MACHINE LEARNING TECHNIQUES:** Regression (Linear, Logarithmic, Polynomial, LassoCV, RidgeCV), Classification (KNN, SVC, Decision Trees, Random Forests, Naive Baysean), NLP, Unsupervised Learning (NMF, LDA), Neural Networks (Convolutional)  
**DATABASES:** SQL, MongoDB  
**DATA SCRAPING:** BeautifulSoup  
**SPOKEN LANGUAGES:** English, French, Chinese  
**IMMUNOLOGY LAB TECHNIQUES:** ELISA Assay, Culture Lysis, Biotinylation, Immunoprecipitation, Western Blot, SDS Page Electrophoresis

### EXPERIENCE

---

#### Metis

Remote

Data Scientist · Sept. 2020 to Dec. 2020

- Metis is an ACCET accredited 12week 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

#### Palo Alto Veterans Affairs Institute of Research

Palo Alto, CA

Research Assistant · Dec. 2018 to Aug. 2019

- Served as a research assistant aiding post-docs on research projects
- Performed image analysis on the movement of foreign white blood cells in the intestinal tract of mice
- Assisted in complex wet laboratory procedures
- Conducted data manipulation and calculation through Microsoft Excel
- Project Manuscript is currently in the review process of the Nature Immunology journal

#### Panera Bread

Lexington, MA

Bakery-Café Associate (Part-Time)

- Balanced the needs of the store with those of customers, focusing on providing a positive customer experience for every individual
- Maintained a cleanly work station

### DATA SCIENCE PROJECTS

---

#### MTA EDA

- Developed a recommendation system for a charity to allocate their limited volunteers based on MTA turnstiles and stations with the greatest usage
- Used historical MTA turnstile records

#### COVID-19 National County Infection Model

- Created a regression model that predicted various American counties' coronavirus infection rates
- Compared Linear, LassoCV, RidgeCV, and Polynomial regression models
- Scraped data from the United States Census, Department of Agriculture, including information on population density, voter distribution, urbanization index, and racial demographics

#### Lung Cancer Risk Prediction

- Developed a classification model that predicted the lung cancer risk of patients using self-reported data
- Compared Logistic Regression, K Nearest Neighbors, Naive Baysean, SVC, Decision Trees to create an optimal classification model
- Created HTML Flask Web app to provide interactive visual for the model

#### Genshin Impact - Twitter Topic Modeling

- Motivated to understand why viral video game Genshin Impact rose to the top Chinese app rankings, yet maintained a very low rating
- Used NMF Topic Modeling to identify the top ten most prominent topics over the course of several weeks on Twitter
- Created LDA plots to visualize spacings of 20 topics on scraped days
- Used snsrape to scrape 120K tweets off twitter with the filter #genshinimpact

#### Ocean Debris Identification with Deep Learning

- Developed a Convolutional Neural Network to differentiate between submerged trash and marine organisms for autonomous trash cleaning devices
- Leveraged Transfer Learning on MobileNet's image detector
- Applied Natural Language Processing on predictions to identify the most commonly encountered objects

### ACTIVITIES

---

#### Fencing ·

US National Team Member, US World Team Member, NCAA Division I Athlete

2014 to 2020

- Represented United States at International World Cups and World Championships
- Peak historical rankings (2017): #2 in National U17 rankings, #3 National U19 rankings, #9 National Senior rankings

### EDUCATION

---

Stanford University · 2018 to Current

BS Bioengineering 2023

Minor Data Science 2023