Li Zhou

E-mail: zh1541@ucsd.edu   **Mobile:** +1 586667916(America)/+86 13675604508
Address: B8-Weihua Xingfucheng, Dingyuan County, Chuzhou, Anhui, 233200, China

Education Background

University of California San Diego   09/2017-06/2021 (Expected)

- **Major:** Data Science, **Minor:** Cognitive Science   GPA: 3.53/4.0
- **Provost Honor** (Spring Qtr 2020, Winter Qtr 2020, Spring Qtr 2019, Fall Qtr 2017)
- **Skills:** Python, Java, SQL, SQL Query, Matlab, R, HTML, CSS, JavaScript, Pytorch

Working Experience

Intern for Recommendation Algorithms, Department of Technology, NetEase Media Group   08/2020-11/2020

- Accessed features based on the LR model from the user side to server, added new ones via Java and Json file.
- Composed feature profiles, dealt with and crossed the original and updated features with completing the training of regression and classification
- Configured samples with matching models through Linux, trained the new feature files and new samples, analyzed the offline AUC and MSE
- Put up the training models on the online AB Testing for the analysis to users' click rate, refresh rate and reading time

Programmer/Data Analyst, UCSD Chinese Computer Community   01/2020-Present

- Participated in the project group of Triple C Kaggle and the related contest, realized the classification of Bengali script via machine learning and the models of high accuracy (96%) analyzed the users' emotions on Twitter
- Cleaned up and classified data, unified the 128*128 format of irregular pictures, self-learned Pytoch and Tensorflow to build the basic model with Pytorch Resnet18 Model with parameter adjustment
- Analyzed the users' emotions expression on twitter, succeeded to won achievement of top 20

Interning Data Analyst, Sany Heavy Industry Co., Ltd.   06/2018-07/2018

- Imported data through Tableau, conducted data analysis and visualize the data to optimize products

Project Experience

CIO, Smart Fridge Student Innovative Startup Project   11/2019-Present

- Managed to solve the problems of food wastage by the combination of algorithms and camera access inside the fridge
- Designed and created a food database, organized the identified data in Google API with data scratch to the food dictionary online, and filtered irrelevant keywords
- Optimized the data cleansing process, co-designed the project web page
- Won the sponsorship of $1,000 from the I-Corp Program of National Science Foundation

AD Analysis on Snapchat   11/2019-12/2019

- Cleaned up the data with Pandas for analyzable factors, analyzed the missing data and determined their types
- Analyzed the factors that influenced the popularity of the advertisement, built basic and models with Pipeline to predict the VV of advertisements

Olympic Data Model   02/2020-03/2020

- Applied SQL to merge Entities(Tables) and create a new schema to conform to 3NF
- Completed a E-R Diagram, and the analysis to Entities, attributes, primary keys, relationship attributes and constraints
- Utilized SQL queries to analyze the conditions of athletes with excellent performance, continuous progress and dominance in every Olympics Games

Score Analysis and Price Expectation on Airbnb (Ranked top 10%)   02/2020-05/2020

- Disposed the data collected from Airbnb in five major cities with Python, fit the price ($/day) on Kaggle Platform
- Established the linear Regression and Multivariate Regression Models to analyze the relationship among the grades, religions, consumptions per person, equipment and the numbers of bedrooms and bathrooms
- Evaluated, screened and handled the features as well as the final MSE, adjusted parameter of the models through Boosting, solved the problem of over-fitting via K-folds

Calibration of Snow Gauge   05/2020-06/2020

- Invented a procedure to calibrate the snow gauge to ensure the precise value of the density reading
- Drew the least square lines for density vs gain and log gain vs density, tested the correlation between two variables with linearity, residual normality, and constant variability of the regression model with R to predict the gain