

CYNTHIA WANG

✉ cynthiawang315@gmail.com

☎ (949)233-8501

📍 Burlingame, CA 94010

in /cynthiawang181128/

🌐 cynthiawang315

EDUCATION

Columbia University in the City of New York

MA Quantitative Methods in Social Sciences (Data Science Focus) 2020

Cumulative GPA: 4.15/4.00

Thesis: Analysis of how comorbid substance use disorders affect suicidal thoughts and behaviors in different age groups with machine learning.

University of California, San Diego

BS Joint Economics & Math 2019

BS Clinical Psychology 2019

Cumulative GPA: 3.92/4.00

SKILLS

PROGRAMMING LANGUAGES

Python

R

MySQL

Git

VISUALIZATION

Tableau

Matplotlib

Seaborn

Plotly

ggplot

Leaflet

MACHINE LEARNING

Regression

Classification

Natural Language Processing

Unsupervised Learning

Neural Networks

Deep Learning

OTHER

HTML

Google Cloud Platform

Flask

Web Scraping

Data Wrangling

Qualtrics

EXPERIENCE

Metis

Data Scientist

San Francisco, CA

06/2020 - 09/2020

- Completed an intensive 12-week project-based data science bootcamp with emphasis on Python programming, data wrangling, machine learning, and communication. See project section for more details.

University of California, San Diego

Research Assistant

San Diego, CA

09/2018 - 06/2019

- Investigated how a disclosure of conflict of interest could affect the effectiveness of defaults as honors thesis.
- Designed survey questions and collected data on Qualtrics; cleaned and analyzed data in JMP and Excel.
- Conducted literature review and facilitated progress on research projects in weekly lab meetings; recruited over 1000 participants for lab studies.

Hangzhou Binjiang Real Estate Group Co

Human Resource Intern

Hangzhou, China

06/2018 - 08/2018

- Integrated staff files for over 1000 employees from 3 company offices in Excel and Human Resource Management System; collaborated with different departments to generate new employee guide.
- Recruited 30 employees and scheduled 400 interviewees monthly; organized recruitment events in 4 cities.

PROJECTS

Classical Music Generation

- Collected 316 classical music midi files and parsed midi files to extract data used in deep learning on Google Cloud Platform.
- Built model architecture with bidirectional GRU and self-attention and generated new songs with the model.

Fairy Tales Topic Modeling and Recommender System

- Scraped, cleaned and transformed the fairy tales using NLTK and spaCy, performed topic modeling with LDA and LSA to explore the topics in fairy tales across continents and developed a recommender system for fairy tales.
- Visualized findings in Tableau and showcased the recommender system in a Flask app.

Titanic Passengers Survival Prediction

- Built classification models(logistic regression, random forest, SVM, XGBoost, KNN) to predict the survival of Titanic passengers with feature engineering to optimize F1-score.
- Created a Flask app that generates survival probability based on user input with the best classification model.

Pet Food Recommendation Score Prediction

- Scraped cat and dog food information such as price and nutrition from Chewy using BeautifulSoup as predictors for pet food popularity.
- Predicted percentages of consumers who recommend the food using linear regression (with L1&L2 regularization) and random forest.

Exploring Dog Ownership in New York City

- Integrated NYC dog licenses dataset with a demographic information API to analyze the dog ownership of New York City residents in R.
- Designed a website to visualize the findings using ggplot, Plotly and Leaflet.