

ISAAC WANG

✉ isaacwang@gmail.com
🌐 medium.com/@isaacwang
☎ (650) 867-8645
📍 Redwood City, CA 94063
in isaac-wang/
🔗 github.com/isaacwang

Data Scientist with a desire to derive actionable insights from data. Excellent analysis, research, and presentation skills developed through work experience and data science projects utilizing a number of tools and methods including, but not limited to: python, natural language processing, SQL/NoSQL, and supervised/unsupervised learning.

SKILLS

TECHNICAL SKILLS

Python
Pandas
SQL
Scikit-learn
Tensorflow
Keras

DATA VISUALIZATION

Tableau
Palantir
Seaborn
Microsoft Excel

LANGUAGES

Fluent Mandarin

EDUCATION

The University of Texas at Austin
BBA Management Information Systems 2018
Certificate in Computer Science
University Honors - Awarded for Outstanding Academic Achievement, 2017

ABOUT ME

Work Eligibility · Eligible to work in the U.S. with no restrictions

Hobbies · Jazz Piano, Triathlons, Tennis, NBA, Hiking, Eating Sushi, Travel

EXPERIENCE

Metis

Data Scientist

San Francisco, NY
June 2020 to Sept. 2020, June 2020 to Sept. 2020

Metis is a 12-week immersive data science program that builds skills in data analysis and machine learning through project development. Completed five-end-to-end. See project section for details.

United Airlines

Data Analyst, Employee Insights

Chicago, IL
June 2019 to Apr. 2020

- Lead the testing and implementation of United Voices, an internal app that collects operational feedback from 40,000 mainline flight attendants
- Built fully interactive department specific (catering, cabin amenities) dashboards in Palantir to help track, analyze and respond to key trending issues
- Utilized Natural Language Processing to identify urgent reports to help prioritize workflow, in addition to performing sentiment analysis on newly added United in-flight products and snacks
- Analyzed user base engagement by applying logistic regression to predict flight attendant churn
- Worked in ServiceNow to manage pending flight attendant reports

United Airlines

Pricing Data Analyst

Chicago, IL
June 2018 to June 2019

- Examined major spikes in refunding passengers and implemented a refundable blackout strategy for newly identified events
- Improved efficiency of marketing process by writing SQL query to more accurately describe seat availability via market
- Constructed a procedure that scans markets for wide gaps between price points to automatically creates mid-level fares, resulting in better segmentation across markets

Quicken Loans

Business Intelligence Data Analyst Intern

Detroit, MI
June 2017 to Aug. 2017

- Collaborated with multiple stakeholders using Agile methodology to create an all-encompassing Tableau dashboard
- Automated an excel report to track QA employee efficiency and reduced reporting time from 1 hour to 5 seconds (see Daniel Wardle's Recommendation further down)
- Extracted data from a variety of different data sources using SQL queries, including joins, aggregate functions, case when statements, temp tables, and subqueries
- Constructed a stored procedure in SQL to automatically populate data into Tableau

Euphoria Music Festival

Data Analyst Intern

Austin, TX
June 2016 to Aug. 2016

- Forecasted festival budget for food, merchandise, and bar vendors, as well as ticket sales
- Estimated artist value for five artists by developing a linear regression model

PROJECTS

Generating Lo-Fi Music with Neural Networks

- Transformed Jazz Piano MIDI files into neural network generated Lo-Fi
- Utilized tensorflow and keras to create jazz piano generation pipeline
- Created a deep learning model with the following layers: Bidirectional LSTM → Attention → LSTM
- Used GCP to train and fit model

Wine Review Text Generator

- Processed and transformed wine review text using Natural Language Processing (NLP)
- Extracted wine topics using LSA, NMF, and LDA
- Visualized topics using pyLDAvis
- Created an internal Markov Chain wine review generator via a Flask application

Classifying Happy vs. Sad Songs

- Applied KNN, Logistic Regression, XGBoost, ADA Boost, and Random Forest models to predict whether a song is happy or not
- Leveraged Spotify API to scrape audio features from multiple playlists
- Used Tableau to visualize distributions of important audio features between happy and sad songs
- Created an interactive happy or sad song predictor via a Flask application
- Deployed the Flask application via Heroku

Predicting NBA points per game

- Developed a linear regression model to predict a total team points scored in a NBA game
- Scraped advanced box score statistics and scores from basketball-reference.com using BeautifulSoup and Selenium
- Discovered that effective field goal percentage and turnover percentage were most indicative of team points a game
- Explored feature and target engineering using polynomial and log transformation

Exploring MTA turnstile data

June 2020 to June 2020

- Experience using pandas to performing exploratory data analysis
- Joined external data sources (Yankee Home Game Schedule, Kiosk Stations) to existing data frames with pandas
- Analyzed passenger traffic changes on Yankee home games vs. Non game days by station
- Visualized data using violin plot via Seaborn and Matplotlib packages