Yiqing (Alyce) Fan

(213)-234-8789 | alycefyq@gmail.com | https://yiqingfan.github.io

EDUCATION

University of Southern California — **Los Angeles, CA**
Master of Science in Analytics, GPA: 3.84 — May 2021
University of California, Los Angeles — **Los Angeles, CA**
Exchange Student, GPA: 3.93 — June 2018
Sichuan International Studies University — **Chongqing, China**
Bachelor of Arts in Business English — June 2019

PROFESSIONAL EXPERIENCE

Southern California Gas Company (SoCalGas, Sempra Energy) — **Los Angeles, CA**
Data Scientist — September 2020 – Present
- Implemented an interactive data entry web application with both front and back ends to assist business team in critical tasks including gas leak detection.
- Implemented Mozilla SOPS for credentials connecting the database for the web application.
- Built a Splunk dashboard to monitor Spark jobs, with key features such as time window pick and developing environments change.
- Developed *Python* ETL scripts via *PySpark* for efficient large-scale data manipulation tasks in SQL server, such joining data across multiple databases and writing records to databases in million scale. Reduced runtime by more than one hour.
- Managed company data by designing schema and migrating across SQL server databases. Leveraging techniques such as partitioning and indexing to optimize data storage and improve query runtime by 20%.
- Utilized Power BI to visualize anomalies, build *Docker* files, and create continuous *Gitlab* integration pipelines.
- Performed code review and implemented automatic code quality check through gitlab-ci.yml.

University of Southern California Integrated Media Systems Center — **Los Angeles, CA**
Student Staff (User-Interface Engineer) Directed by Dr. Luciano Nocera — January 2021 – May 2021
- Developed a dashboard in *Vue.js* for freight modeling research that displays sensors, truck states, and traffic.
- Utilized *Mapbox* and bootstrap table to help researchers improve the algorithm, and improve client understanding.
- Visualized modeling results by displaying the estimation matrix and ground truth matrix with heat maps.

Jaguar Land Rover — **Shanghai, China**
Digital Marketing Analyst Intern — March 2019 – June 2019
- Drafted KPIs for convertion funnel and consumer profiles to track customer behavior.
- Designed the *Brand App*, modified user story descriptions based on data analytics, and worked with multiple departments.
- Reported selling results monthly via Google Analytics and assisted the senior manager in adjusting market strategies.

China Unicom — **Beijing, China**
Big Data Analyst Intern — June 2018 – August 2018
- Led a team of four members to monitor and analyze the 2018 World Cup buzz by employing a *Python*-based web crawler and sentiment analysis tool for automated data mining to classify online news.
- Implemented an HTML5-based front-end presentation tool to publish analysis results to the WeChat social media platform.
- Built a positioning model in *Python* to predict customers' locations by applying the random forest algorithm and training the dataset. The model reached an accuracy of 96.57% against test data.

GRADUATE STUDENT PROJECTS

Security Problem Solution for RIOgaleao Airport | *Python* — February 2021 - Present
- Performed data analytics to separate passengers and staff based on Mac addresses data counts of connection times.
- Displayed passenger/staff hourly/daily distribution/density patterns via interactive heat map and gridding technology.
- Built a ARIMA time series model to enable risk level labeling for each grid in the airport.

Storytelling for Netflix | *JavaScript* — August 2020 - December 2020
- Performed data preparation with a *Kaggle* dataset and *Python*.
- Designed responsive and interactive D3 Charts, a D3 map, and a *Mapbox* map to visualize Netflix producers and audiences.
- Built a website using the *Vue.js* framework, and presented charts according to a storyline.

SKILLS AND CERTIFICATIONS

- **Languages:** Mandarin (native), English (fluent, literate), French (rudimentary).
- **Programming Languages:** *Python*, *PySpark*, *R*, *SQL (SQL Server Management Studio)*, *JavaScript (D3.js, Vue.js, React.js)*.
- **Tool Kits:** *GitLab/GitHub*, *Docker*.
- **Other:** *Splunk, Google Analytics, Power BI, Tableau, Figma, Microsoft Excel.*
- **Certifications:** IBM Applied Data Science Specialization (Issued June 2020)