# Chang (Cici) Chang

(213)706-7927 | cicichan@usc.edu | https://github.com/cicichang11

## EDUCATION

**University of Southern California** **Los Angeles, CA**
*B.S. in Industrial and System Engineering- Information Systems Focus* Aug 2019 – May 2023
*M.S. in Applied Data Science* Aug 2022 – May 2024

- **Minor:** Applied Analytics; Mobile App Development; Computer Science
- **Undergrad Overall GPA:** 3.8/4.0; **Graduate GPA:** 4.0/4.0
- **Honors:** Dean's List at USC Viterbi School of Engineering, USC AAA Scholarship, Center of Undergraduate Research Fellow
- **Courses:** Machine Learning, Database Management, Data Structure, Enterprise Data Analytics, Operations Research

## TECHNICAL SKILLS

- **Programming:** Python (Pandas, Scikit-Learn, PyTorch, Matplotlib), SQL, JavaScript, HTML, CSS, Swift, Java
- **Infra, Database & Tools:** Power BI (DAX, Power Query), Tableau, Excel, PowerPoint
- **Statistical Methods:** ANOVA t-tests, A/B testing, Linear/Logistic Regression, Decision Trees, Time Series Analysis

## PROFESSIONAL EXPERIENCE

**WANB (Startup)** **Los Angeles, CA**
*Data Scientist Intern – Data Automation & Warehousing* May 2023 – Present

- Developed Flask web applications in Python to analyze bidding data, consolidate vendor quotes, compare pricing, and generate interactive reports with evaluation metrics to streamline bid analysis
- Modeled a TensorFlow neural network to optimize inventory routing and minimize shipping costs, compiled with a cost & on-time delivery penalty loss function and Adam optimizer, improving on-time delivery rates by 20%

**Amgen** **Thousand Oaks, CA**
*Data Scientist Intern – Digital Workplace Experience* May 2022 – Aug 2022

- Developed a comprehensive data integration process, utilizing Microsoft Viva and Glint to measure employees' performance and provide in-depth insights into 12 categories of employee behavior and sentiment
- Built relational databases with 60 data tables using Power BI and Power Query to automate survey results reporting processes
- Identified 8 KPIs, created 16 pages Power BI report, and developed 6 key metric correlations to drive targeted actions
- Developed and maintained data pipelines using Spark SQL in Databricks to automate the ETL process of the data lake

**USC Viterbi School of Engineering** **Los Angeles, CA**
*Data Science Research Assistant* Sep 2021 – Present

- Developed an age-sensitive time-varying SEIRD (Susceptible, Exposed, Infectious, Recovered, and Deceased) model using Python to accurately portray dynamic trends of COVID-19 cases and deaths for 50 states in the USA with RMSE lower than 4%
- Optimized vaccine allocation strategies for the epidemic by leveraging optimization algorithms, resulting in achieving a significant reduction of about 1.9 million accumulated cases and deaths in the USA
- Leveraged logistic function to model the infection progression to visualize the fitting results in S-shaped curves

**JD.com** **Beijing, China**
*Data Scientist Intern – To-B Payment Product Team* Apr 2021 – July 2021

- Designed MySQL relational databases for customer feedback submission system to improve data consistency and performed data mining and classification using Python on the data of users' problems
- Leveraged JD Cloud Computing to automate ETL processes and report on campaign result datasets, improving efficiency by 40%
- Performed A/B testing, prioritized features, and drove new UI design that reduced 36% of users' complaint rate
- Defined KPIs and developed Power BI dashboards to monitor key metrics, providing insights into business performance

**Bain & Company** **Shanghai, China**
*Data Analyst Intern – Management Consulting Part-time Assistant* May 2020 – Oct 2020

- Conducted market research and assessed competitive landscape of the medical aesthetic industry to draft 5 market entry strategies
- Used SQL to analyze obtained datasets and built data visualizations of key insights on client-faced presentation slides

## PROJECT EXPERIENCE

**Detect Machine Failure from IoT Sensors with a SQL Pipeline** **Los Angeles, CA**
*Methodologies: Machine Learning, Classification, Streamlit* July 2023

- Developed an end-to-end SQL pipeline to detect early signs of industrial machine failure from IoT sensor data, that increased machine failure prediction accuracy by 15% compared to previous methods
- Performed exploratory data analysis, preprocessing, and feature engineering, including PCA for 50% dimensionality reduction
- Improved model F1 score from 0.82 to 0.94 through hyperparameter tuning and cross-validation and deployed top-performing model SGD, Ridge, Random Forest, and Gradient Boosting into production
- Automated alerts when failure risk passed the 15% threshold, enabling timely maintenance scheduling before critical breakdowns
- Implemented model retraining on new data for continuous improvement