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
- **Graduate GPA:** 4.0/4.0; **Undergrad Overall GPA:** 3.8/4.0
- **Courses:** Machine Learning, Database Management, Data Structure, Enterprise Data Analytics, Operations Research

TECHNICAL SKILLS

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

PROFESSIONAL EXPERIENCE

WanB Express — **Los Angeles, CA**
Data Engineer 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%
- Built a generative AI Chatbox to answer customer service questions using OpenAI API and AWS Cloud Computing
- Developed an automated machine learning web app via PyCaret and Streamlit, improving development efficiency by 40%

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
- Built relational databases with 60 data tables using Power BI and Power Query to automate survey results reporting processes
- 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 to reduce 1.9 million accumulated cases and deaths, leveraging optimization algorithms
- 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

- Built a natural language processing model in Python using scikit-learn for the customer feedback submission system to analyze service reviews and classify sentiment, using techniques like vectorization, TF-IDF, and SVM
- 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

Bain & Company — **Shanghai, China**
Data Analyst Intern – Management Consulting Part-time Assistant — May 2021 – Oct 2021

- 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

Chinasoft International — **Beijing, China**
Software and Big Data Engineer Intern — May 2020 – July 2020

- Developed Linux Docker containers for various software applications, including MySQL and PostgreSQL, created Docker files, used Git tags to mark releases of containers, installed software dependencies, and configured the container environment
- Developed a web application backend using Flask and PostgreSQL, set up user authentication, REST API endpoints, and database

PROJECT EXPERIENCE

Movie Recommender System Web App — **Los Angeles, CA**

- Developed and deployed a movie recommender web application using Streamlit that provides users with movie suggestions based on cosine similarity between movie metadata vectors
- Engineered an item-item collaborative filtering algorithm that compares movie plots, casts, keywords, and genres to build a similarity matrix using scikit-learn; implemented model training and saving pipeline
- Containerized services using Docker for reproducible model training and serving; deployed to Heroku with CI/CD and integrated with Streamlit front-end application

Customer Churn Prediction Web App — **Los Angeles, CA**

- Developed a machine learning pipeline to predict customer churn using scikit-learn, and performed exploratory data analysis and feature engineering to improve model AUC by 15%
- Built and optimized classification models including SGD, Ridge, Random Forest, Gradient Boosting, and XGBoost
- Created a web application with Streamlit to allow dynamic customer churn predictions based on input features