

SIYUAN FENG

siyuanfeng09@gmail.com | (585) 851-2578 | <https://www.linkedin.com/in/siyuanfeng/> | <https://siyuan09.github.io>

SKILLS

Programming Languages: R (ggplot2, caret, dplyr, rpart), Python (Numpy, Pandas, Scipy, Matplotlib, Seaborn, Scikit-Learn, Plotly).

Analytical Tools: SQL, NoSQL, Tableau, PowerBI, Mode, Advanced Excel, AWS, Google Analytics, Bloomberg Terminal.

Statistical Learning: Regression (Linear, Logistic), Decision Tree, Random Forests, Naive Bayes, SVMs, KNN, Adaboost, Clustering.

Soft Skills: Business Analytics, Problem-solving, Professional communications, Teamwork, Leadership.

PROFESSIONAL EXPERIENCE

Ibotta, Inc. - Denver, CO

Feb. 2021 - Present

Associate Data Scientist

- Led the development and implementation of Bayesian regression media mix model using 5+ years of data to optimize marketing spend for user acquisition, contributed to a 10% increase in installs.
- Introduced new user segmentation using K-means clustering and decision trees based on behavioral characteristics, designed and tested customized activation tactics for different groups, and efficiently raised the activation rate by 6%.
- Maintained and automated machine learning models on Airflow, adjusted the weights of the objective function and population size based on budget pacing, and tracked the model performance by building an interactive dashboard using SQL and Python; saved 60% of labor time and reduced \$1.4M cost.
- Designed and implemented disparate experiments by understanding business strategy and underlying compliance risk, conducting cohort analysis, deciding target audience, timeline, and group size, and building statistical models such as multivariate regression and logistic regression; validated different activation tactics and increased the activation rate by 5%.
- Closely collaborated with and supported the growth marketing team by providing ad-hoc root-cause analysis, hypothesis testing, and other analytical tasks; made business recommendations with effective presentations of findings.

Vee Ventures - New York, NY

July 2020 - Sept. 2020

A Management consulting firm providing organization design, executive search, and recruitment for Fortune 500 companies.

Data Scientist Intern

- Conducted analysis of the recruiting process such as application drop-off rate and reasons for losing candidates; optimized the process by adjusting screening steps, time intervals, and focus; decreased the drop-off rate by 2%.
- Clustered candidates into different potential groups by K-means clustering based on competencies scores and interview scores; improved the candidate pool and search efficiency.
- Managed the data by extracting and transforming 120k+ records of 60+ tables from AWS Redshift; conducted ad-hoc analysis using SQL and Python according to stakeholders; developed a real-time ETL pipeline for data processing; maintained Python code by Git.
- Designed a daily-updated interactive dashboard using PowerBI; enhanced the dashboard via high-tempo testing and feedback cycles; put the dashboard into production for daily reporting, and effectively saved 30% of the time in daily meetings.

L'Oréal China - Shanghai, China

Sept. 2018 - Mar. 2019

Data Analyst Intern

- Contributed to the promotion of 10+ products by analyzing the products' attributes, prices, and campaign strategies; identified potential competitors and market trends by comparing sales performance and analyzing promotion trends in brand posts.
 - Web scraped 100k+ rows of online posts of customers; leveraged NLP techniques to analyze the content, and performed sentiment analysis using Python to determine customers' pain points; provided insights for marketing strategies; contributed to 20k+ shares on social media and 10k+ unit sales of the new product in the first month.
 - Extracted and analyzed 200k+ sales data using SQL every week; designed the visualizations to evaluate sales performance.
-

PROJECT EXPERIENCE

Wine Retailer Email Campaign Experiment Analysis - Rochester, NY

Jan. 2020

- Conducted A/B testing and verified the impact of sending emails on driving purchases by analyzing the average causal effect in R.
- Identified the target customers by the slice and dice analysis; built the causal forest model to predict individual purchase lift, and sent emails to 56% of customers who generate positive profit; raised the profit by \$14,000.

HELOC Risk Prediction Machine Learning Project - Rochester, NY

Dec. 2019

- Performed exploratory data analysis (EDA) on 10k+ client records; built models such as Logistic Regression, SVMs, Random Forests, Adaboost, and Neural Networks as well as tuned parameters; achieved 72% prediction accuracy on test data in Python.
 - Designed an interactive interface using Streamlit in Python, and provided explanation of feature contribution using LIME in Python.
-

EDUCATION

Simon Business School, University of Rochester - Rochester, NY

Dec. 2020

Master of Science in Marketing Analytics (STEM)

Fudan University - Shanghai, China

June 2019

Bachelor of Arts - Advertising