Wenjing (Cali) Li

734-882-5318 | wenjingli199707@gmail.com | GitHub.com/cali-li | www.LinkedIn.com/in/cali-li/

OBJECTIVE

Professional Data Scientist skilled at analyzing customers and delivering data-driven recommendations to business partners. Passionate and experienced in executing modeling projects, ensuring their accuracy, and elevating the client experience. Proficient in bridging business demands with technical solutions to drive successful completion of mission-critical projects.

EXPERIENCE

Senior Data Scientist, Ford Credit, MI
Jan 2023 – Current

- Led a team of three in developing and implementing statistical and XgBoost models to predict customers' future purchases, achieving a 6% accuracy increase. This enhanced the effectiveness of personalized offers and messages.
- Designed and monitored a CLTV (customer lifetime value) model integrated with statistical and time series models (SARIMA) to forecast the product portfolio, contributing to the establishment of business objectives and priorities.
- Attained proficiency in SAS through self-learning, and improved efficiency by optimizing a cloud-based SAS script, reducing the runtime from 360 minutes to 30 minutes.

Data Scientist, Ford Credit
Jun 2021 – Jan 2023

- Conducted in-depth analysis of the retail customer population, resulting in the identification and characterization of a previously undiscovered customer sub-segment. This insight has been integrated into new digital products.
- Analyzed 6 experiments using AB testing to introduce new digital features, resulting in a savings of 1.2 million.
- Designed and automated vehicle payoff performance reports to support employee recruitment and resource allocation.
- Analyzed clickstream data via Logistic Regression to identify a clicks-to-sales correlation for 20K clicks.
- Utilized various platforms including Hadoop, Teradata, PC, Linux, and GCP to compile data and investigate data issues.

Data Analysis Assistant, University of Michigan, MI
Sep 2020 – May 2021

- Devised PySpark-based filtering methods for COVID-related tweets, reducing computation time from 480 to 50 minutes.
- Performed exploratory data analysis on 50 billion tweets with PySpark to examine COVID's impact on people's lives.

Visualization Consultant, University of Michigan
Jul 2020 – May 2021

- Constructed interactive visualizations for the SCF and PSID datasets measuring financial standing and race disparity.
- Hosted a real-time, communicative household financial situation visualization dashboard on a website using JavaScript.

Data Visualization Consultant, University of Michigan
Oct 2019 – May 2020

- Built Next Word Prediction model using NLP methods; created a Python-based web application for its implementation.
- Preprocessed math publication data from 1840 to 2019 and produced an interactive network visualization using D3.js.

Key Account Intern, LinkedIn, China
Dec 2018 – May 2019

- Translated marketing behavior analyses into Quarterly Business Reports, offering insightful marketing proposals.
- Monitored 8,000 monthly recruitment activities, delivering LinkedIn Solution Usage feedback to clients.
- Contributed to training programs for clients and built connections between recruiting clients and potential candidates.

EDUCATION

M.S. in Applied Statistics, 2021
University of Michigan - Ann Arbor

B.Sc. in Mathematical Statistics and (minor) B.A. in Finance, 2019
Sichuan University

SKILLS

- Python, SAS (EG, EM & Viya), SQL
- Teradata, Hadoop, GCP
- Time Series modeling
- Tableau, Qlik, Alteryx
- Statistical Modeling
- Data Visualization
- Numpy, Pandas, TensorFlow, PyTorch
- Machine Learning (NLP, XgBoost)
- LaTeX, Git, Linux