Liang Zhao [in]

Open to relocate | (805)-689-6717 | lz2871@columbia.edu | github:liang0711 | Tableau

EDUCATION

Columbia University   Expected Dec. 2023
Master of Science in Statistics (STEM)

University of California, Santa Barbara   09/2018 – 06/2022
Bachelor of Science in Actuarial Science

PROFESSIONAL EXPERIENCE

Columbia University   New York, NY
Research Assistant (Part-time)   06/2023 – 07/2023
- Web scraped 13k+ rows of unemployment benefit appeals related from NY official government site using Python crawler (selenium & Beautiful soup) to select related case information from PDF files, which increase working efficiency 95%
- Unified data for NY state from 25K+ rows of monthly Excel files using Python and automated the transform and load process
- Optimized data management based on the type of court, appellant and won cases to identify the appeals pattern over years in Tableau Dashboard and delivered data insights, which resulted in searching time reduction from 3 hour to 1 hour

China International Capital Corporation, Wealth Management   Shenzhen, China
Project – Data analyst Intern   s 06/2021 – 08/2021
- Developed a data-based mobile app for financial product releasing purposes and initiated a customer/marketing/strategy analytics framework for product development and optimization purpose
- Participated in the new feature page prototype design of mobile APP to improve the user experience, especially upgrading, and denoting the page design in computer/Mobile version
- Visualized bi-weekly reports by using in-house BI dashboard on customer credit history and financial summaries to guide senior leader decisions on optimizing KPI based on metrics click-through-rate and retention rate for building UAT test/time series analysis
- Collaborated with three technology and operation teams to ensure data reporting and business requirements by using agile methodology to ensure successful feature launching

Tencent Inc.   Shenzhen, China
Business Intelligence Intern   10/2020 – 12/2020
- Developed 10+ daily user reports using Python and SQL, facilitated operation team and stakeholders in timely customer management for 400K+ data rows and inventory management for 3000+ SKUs
- Conducted outlier analysis by analyzing duplicated (user_ID, date, amount, etc.) in Python, identifying abnormal transaction behavior, which generated 100+ risk activities with risk level and resulted in decreasing company's budget loss 20%
- Designed reports for customer segmentation for major user groups which increased total sales by 36% in next marketing campaign to recommend 10 top-ranking products from dashboard
- Proposed 3+ new product lines and identified marketing campaign opportunities to lead new GMV from 80% previous users

PROJECT EXPERIENCE

New York Housing Price Dashboard | Tableau Analytic, Project   06/2023 – 06/2023
- Gathered and cleaned housing market 2k+ data from API sources, developed ETL process to ensure data accuracy and consistency
- Created interactive Tableau dashboards that showcased essential KPIs such as average home prices, market status levels etc. Developed dynamic filters and parameter within Tableau to allow users to explore data based on location/ price range/home type

Google Play Store Sentiment Analysis | NLP Course Project   03/2023 –05/2023
- Performed EDA on 50k rows of Google Play Store App reviews of Android market and business ratings using Numpy and Pandas
- Vectorized, removed punctuation and stopwords, stemmed and lemmatized the bag of words using nltk package in Python
- Applied and tuned Logistic regression, SVM, Random Forest to predict the sentiment, which achieved an accuracy of 90%

Ecommerce Product Arrival Analysis | ML Modeling Project   03/2023 – 05/2023
- Conducted feature engineering on 4k+ rows of data and selected 5 features by performing variable transformation and imputation
- Applied and tuned hyperparameters on Decision Tree/XGBoost model to predict the arrival with accuracy with 70%

TECHNICAL SKILLS

Technical Skills: Python (Pandas, NumPy, Scikit-Learn), SAS, SQL(MySQL), R (ggplot2), Spark, Google Cloud Platform
Software: Excel (VLOOKUP, Pivot Table), Tableau, Power-BI, Word, PowerPoint, Google Analytics, AWS (Sage-Maker)
Data Engineering: Data Extraction, Manipulation and Database Management, AWS