Yi-Li Chen

+1 (323)336-2382 | leochenusc@gmail.com | https://www.linkedin.com/in/yi-li-chen-leo/

PROFESSIONAL SUMMARY

Detail-oriented professional with 5+ years of hands-on data analytics experience in SQL, Python, R and Tableau, translating business requirements into solutions, and collaborating with cross-functional teams to drive data-driven decision-making.

WORK EXPERIENCE

SmartBI LLC (getBIsmart.com), Irvine, California

Lead Analyst   *Jul 2022 – Present*

Data Analyst   *Jul 2021 – Jul 2022*

- Analytics Consultant for entertainment companies such as Sony Pictures, Crunchyroll, Scopely and Laguna Games
- Optimized ETL processes for multiple data sources and **90M+** user accounts and increase efficiency by **30%**
- Utilized the Snowflake cloud data warehouse and AWS S3 bucket to seamlessly integrate data from diverse sources
- Designed KPIs and performed cohort analysis on in-product data, increasing conversion rate by **14%**
- Implemented and evaluated **A/B tests** with product teams to fine-tune features and increase retention rate **~10%**
- Built models in **Python** to forecast CLTV and classified users using XGBoost for campaign performance
- Automated ad-hoc analytic reports in **SQL** to provide marketing performance insights for monthly ROAS
- Provided insights with interactive dashboards in **Tableau** to synthesize complex analytical tasks into data-driven stories
- Collaborated with Marketing, Product and Engineering teams and provided expertise for applied analytics team
- Assist with new hires training process and document on-board program setup and learning manuals for new hires

Academia Sinica, Taipei, Taiwan   *Apr 2019 – Jul 2019*

Research Assistant (Data Analysis)

- Preprocessed the data from noisy wireless signal by data cleaning, integration and transformation
- Conducted **Python** and **SQL** analyses on over **1-million** data from noisy wireless signal
- Provided quantitative, visual explanations and critical features for machine learning process and prediction
- Utilized visualization techniques such as PCA and FastMap to project the high-dimensional features to the 2D space

EDUCATION

University of Southern California (USC), Los Angeles, CA   *Aug 2019 – May 2021*

Masters in Analytics

Relevant Coursework: Machine Learning for Data Science, Data Management, Data Mining, Optimization Methods for Analytics, Text Analytics, Integrative Analytics, Predictive Analytics

PROFESSIONAL SKILLS

- Programming: SQL, Python (*scikit-learn, pandas, numpy*), R (*tidyverse, tree, rpart*)
- Toolbox: Jupyter, PyCharm, R Studio, Git, MySQL, pgAdmin, Snowflake, DBeaver, AWS
- Visualization: MS Excel, Tableau, tidyverse, matplotlib, seaborn
- Expertise: EDA, Statistical Modeling, A/B Testing, Predictive Analytics Modeling, Data Mining, Data Visualization

PROJECTS

Recipe Recommendation System (Python, Web scraping)   *Spring 2020*

- Built a recipe recommendation system that can help users to customize their meal by desired preference in Python
- Scraped several hundred-thousands of recipes and reviews from websites with Selenium and BS4
- Optimized user-based weekly meal planning system using Pyomo library with **~90%** level of satisfaction

NLP-based Comedian Recommender System (Python, Clustering, Classification, Hypothesis Testing)   *Fall 2020*

- Recommend a stand-up comedian from given look-alike comedians based on scripts content with Machine Learning
- Applied LDA for content-based filtering on transcripts and vectorized comedians with GMM
- Applied sentiment analysis and matrix completion on the scraped comments to build users matrix
- Predicted user's preference using KNN with combined matrix and improved its accuracy rate by nearly **10%**

Airbnb Data Mining (R, Exploratory Analysis)   *Fall 2020*

- Analyzed and visualized **hundred-thousands** of crime and housing data using tidyverse/*tree*/rpart libraries in R
- The accuracy of predicting prices of Airbnb with cluster analysis and decision tree is greater than **85%**

Hospital Performance Analysis and Visualization (SQL, Excel, Tableau)   *Spring 2021*

- Collaborated with Keck hospital and supported ad-hoc analytics using SQL/Excel with interactive charts in Tableau
- Visualized the relationship between the healthcare KPIs like wait time and equipment utilization and seasonality

Mobile Game A/B Testing (Python, A/B Testing)   *Spring 2023*

- Conducted an A/B test to assess the impact of varying gate levels in a mobile game on player retention
- Leveraged pandas and plot function with hypothesis testing and bootstrapping, yielding valuable statistical insights.