Yi-Li Chen

✉ leochenusc@gmail.com   ☐ 3233362382   📍 Elk Grove, CA   ☐ public.tableau.com/app/profile/yi.li.chen

PROFESSIONAL SUMMARY

Detail-oriented data analytics professional with 5+ years of experience in SQL, Python, R, and Tableau, skilled in translating business requirements into solutions, collaborating with cross-functional teams, and driving data-driven solutions to improve business outcomes.

WORK EXPERIENCE

SmartBI LLC (getBIsmart.com)   *Irvine, CA*
Lead Analyst   *Jul 2022 – Present*
- Led client meetings to identify needs and define strategies, converting stakeholder requirements into actionable features
- Streamlined data processing workflows and supported daily operations with ad-hoc analyses using **Python** and **SQL**
- Designed KPIs and conducted cohort analysis on in-product data, resulting in ~**15%** increase in retention rate
- Implemented and assessed A/B tests with product teams to refine features, leading to a **10%** increase in conversion rate
- Created loyalty strategies through customer segmentation, achieving a 40% rise in engagement and 15% revenue growth
- Investigated suspicious account activities, analyzing transaction patterns and user behavior to identify potential fraud
- Assisted with new hire training and documented onboarding processes and learning materials

Data Analyst   *Jul 2021 – Jul 2022*
- Served as an analytics consultant for entertainment companies such as Sony Pictures, Crunchyroll and Scopely
- Optimized ETL processes for multiple data sources and **90M+** user accounts and increase efficiency by **30%**
- Utilized **Snowflake** and **AWS S3** to seamlessly integrate data from diverse sources, ensuring consistency and reliability
- Provided insights with interactive dashboards in **Tableau** to synthesize complex analytical tasks into data-driven stories
- Built predictive models in **Python** to forecast CLTV and utilized XGBoost for user segmentation in targeted campaigns
- Collaborated with Sales, Marketing, Product and Engineering teams and provided expertise for applied analytics team

Academia Sinica   *Taipei, Taiwan*
Research Assistant (Data Analysis)   *Apr 2019 – Jul 2019*
- Preprocessed the data from noisy wireless signal by data cleaning, integration and transformation
- Conducted **Python** and **SQL** analyses on over **1-million** data points 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*
Masters in Analytics   *Aug 2019 – May 2021*
Relevant Coursework: Machine Learning for Data Science, Data Management, Data Mining, Predictive Analytics

PROFESSIONAL SKILLS

- Programming: Python (*Scikit-learn, SciPy, Pandas, NumPy, Seaborn, Matplotlib*), R (*tidyverse, tree, rpart*)
- Toolbox: Jupyter, PyCharm, R Studio, MySQL, PostgreSQL, Snowflake, AWS (*S3, Lambda, Redshift, SageMaker*)
- Visualization: MS Excel, Tableau, Google Data Studio, QuickSight, Power BI
- Expertise: EDA, Statistical Modeling, A/B Testing, Predictive Analytics Modeling, Data Mining, Data Visualization

PROJECTS

Recipe Recommendation System (Python, Web scraping)   *Spring 2020*
- Constructed a recipe recommendation system for meal customization, scraping recipes using Selenium and BS4
- Optimized user-based weekly meal planning system with Pyomo library, achieving ~**90%** satisfaction

NLP-based comedian recommender system (Python, NLTK, TextBlob, Hypothesis Testing)   *Fall 2020*
- Recommended stand-up comedians based on script content using Machine Learning
- Implemented content-based filtering analyses on transcripts and vectorized comedians with GMM
- Applied sentiment analysis and matrix completion on comments, enhancing user preference prediction accuracy by **10%**

Airbnb Data Mining (R)   *Fall 2020*
- Analyzed and visualized **hundreds of thousands** of crime and housing data using tidyverse, ggplot, and rpart libraries
- Achieved over **85%** accuracy in predicting Airbnb prices with cluster analysis and decision trees

Hospital Performance Analysis and Visualization (SQL, Execl, Tableau)   *Spring 2021*
- Collaborated with Keck hospital, supporting ad-hoc analytics using SQL/Excel and interactive charts in Tableau
- Visualized relationships between healthcare KPIs, such as 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