Yi-Li Chen

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

WORK EXPERIENCE

SmartBI LLC (getBIsmart.com), Irvine, California   *Jul 2021 – Present*
Data Analyst
- Developed and maintained ETL process with data integration and implemented stored procedures
- Defined mobile game KPIs and performed cohort analysis with segmentation and improved conversion rate by 14%
- Managed ad-hoc analytics and built reports to provide marketing performance and increase monthly ROAS
- Implemented and evaluated A/B tests with product teams to fine-tune features and increase retention rate
- Built model in Python to predict CLTV and classified users in machine learning for campaign performance
- Provided business insights with interactive dashboards to synthesize complex analytical tasks into data-driven stories
- Collaborated with Marketing, Product and Engineering teams and provided expertise for applied analytics team

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 several visualization techniques to project the high-dimensional feature space 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

National Central University (NCU), Taoyuan, Taiwan   *Sept 2017 – Nov 2018*
Masters in Mechanical Engineering

PROFESSIONAL SKILLS

- Programming: SQL, Python (*scikit-learn, pandas, numpy*), R (*dplyr, tidyverse*), JavaScript
- Computer Skills: Jupyter, PyCharm, R Studio, Git, MySQL, PostgreSQL, Snowflake, DBeaver, AWS
- Visualization: MS Excel, Tableau, ggplot2, Matplotlib, Seaborn
- Expertise: NLP, Text Mining (*nltk*), Web Scraping, Regression, Clustering, Classification, Exploratory Data Analysis, Statistical Modeling, A/B Testing, Time-Series Forecasting, Data Mining, Data Visualization

PROJECTS

A linked knowledge base of crime data (Python, pgAdmin, PostgreSQL, Tableau)   *Fall 2019*
- Built ETL flow and integrated crime dataset with pgAdmin and analyzed over 2 million entries in Python
- Created data-driven solution and visualized the result by using Seaborn and Matplotlib

Recipe Recommendation System (Python, Web scraping)   *Spring 2020*
- Built a recipe recommendation system that can help users to customize their meal by desired preference
- Scraped several hundred-thousands of recipes and reviews from websites with Selenium and BS4
- Modeled an optimization problem for meal planning system with Pyomo library after collaborative filtering
- Provided an optimized user-based weekly meal plan with 90% level of satisfaction

NLP-based comedian recommender system (Python, NLTK, TextBlob, Hypothesis Testing) *Fall 2020*
- Recommend a stand-up comedian from given look-alike comedians based on scripts content
- 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 in ML with combined matrix and improved its accuracy rate by 6%

Airbnb Data Mining (R)   *Fall 2020*
- Incorporated several hundred-thousands of crime data and recommend housing locations for leasing
- Analyzed and visualized the cleaned data by using tidyverse/ggplot/ggmap/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, Execl, Tableau)   *Spring 2021*
- Conducted research with Keck hospital and improve management with KPI in healthcare
- Composed ad-hoc analytic reports by using SQL with interactive charts via excel and Tableau
- Visualized the relationship between the healthcare KPIs like wait time and equipment utilization and seasonality