

# Xiao (Seanny) Song

☎ 2405641321 ✉ [songxiao@umd.edu](mailto:songxiao@umd.edu) 🗣 songxxiao 🌐 [xsong.ltd](https://xsong.ltd) in seannysong

## Education

Robert H. Smith School of Business, University of Maryland, College Park, MD 2021~Present  
Master of Business Analytics

East China Normal University, Shanghai 2016~2020  
Bachelor of Sociology

## Work Experience

**FancyDigital** 2020-09~2021-04  
*Data Analyst(Full-Time)*

- Use distributed OLAP databases such as Apache Doris/TiDB/Clickhouse to write BI data dashboards. Use a variety of visualization methods to show usage and revenue of system. Count consumption of advertising in multiple dimensions, such as time and user dimensions
- Perform user churn analysis on company's SaaS products, provide user service suggestions to operators, and ultimately reduce user churn rate to 10%
- Analyzed error log during product use and assist product manager to optimize product design and user experience. Successfully reduced bug rate to below 5%

**Zhongnan University of Economics and Law** 2020-02~2020-04  
*Data Scientist Internship*

- Used Xgboost, RandomForest, LightGBM and or algorithms to classify (multiclass) legal text data
- Preformed word frequency method to construct feature matrix, and cross-validation training model (`sklearn`) is used to obtain cross-validation accuracy of 0.75

## Awards and Honors

**Kaggle M5 Forecasting - Accuracy** 103rd/5558 Top2% Silver Medal 2020  
*Estimate unit sales of Walmart retail goods*

•

## Technical Skills

- **Database / Data Warehouse** Good at writing and optimization of sql commands based on MySQL, SQL Server, etc. Familiar with advanced methods such as window functions
- **Data Analytics & Statistics** Familiar with data manipulation and statistical analysis based on Python and R language, including use of software packages such as `numpy`, `pandas` and `data.table`, `tidyverse`
- **Machine Learning** Familiar with ory and application of traditional machine learning methods, including cross-validation training and parameter adjustment based on `sklearn`'s Logistic Regression, SVM, Xgboost and or models
- **Data visualization** Proficient in using Tableau, `ggplot2`, `plotly` and or software packages to visualize discrete data and continuous data, and be good at making dashboards based on business logic
- **Web Crawler** Familiar with basic knowledge of html/css/json, can use Selenium, bs4 and or tools in Python to crawl data presented in webpage and write it into database for storage
- **Other Skills**  
Linux/Shell, SPSS, Stata, Git,  $\text{\LaTeX}$ , HTML/CSS