Xiao (Seanny) Song

songxxiao ▪ 2405641321 ▪ songxiao@umd.edu ▪ xsong.ltd

Education

Robert H. Smith School of Business, University of Maryland, College Park, MD — 2021~2023
Master of Business Analytics

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

Professional 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 the usage and revenue of the system. Count the consumption of advertising in multiple dimensions, such as time and user dimensions.

- Perform user churn analysis on the company's SaaS products, provide user service suggestions to operators, and ultimately reduce the user churn rate to 10%.

- Analyze the error log during product use and assist the product manager to optimize product design and user experience. Successfully reduced the bug rate to below 5%.

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

Use Xgboost, RandomForest, LightGBM and other algorithms to classify (multiclass) legal text data. The word frequency method is used to construct the feature matrix, and the cross-validation training model (`sklearn`) is used to obtain the cross-validation accuracy of 0.75. I write a program to make predictions on new data, so that the prediction results can be applied to any new data set.

Awards and Honors

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

Kaggle M5 Forecasting - Uncertainty   18th/909 Top2% Silver Medal — 2020
Estimate the uncertainty distribution of Walmart unit sales

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 the use of software packages such as `numpy`, `pandas` and `data.table`, `tidyverse`.

Machine Learning Familiar with the theory and application of traditional machine learning methods, including cross-validation training and parameter adjustment based on `sklearn`'s Logistic Regression, SVM, Xgboost and other models.

Data visualization Proficient in using Tableau, `ggplot2`, `plotly` and other software packages to visualize discrete data and continuous data, and be good at making dashboards based on business logic.

Web Crawler Familiar with the basic knowledge of html/css/json, can use Selenium, bs4 and other tools in Python to crawl the data presented in the webpage and write it into the database for storage.

Other Skills
Linux/Shell, SPSS, Stata, Git, LaTeX, HTML/CSS