Xiao (Seanny) Song

(240)564-1321 songxiao@umd.edu songxxiao xsong.ltd seannysong

EDUCATION

Robert H. Smith School of Business, University of Maryland, College Park, MD — Aug 2021 ~ Dec 2022
Master of Science in Business Analytics (STEM), Terrapin Scholarship

East China Normal University, Shanghai, China — Sep 2016 ~ Jun 2020
Bachelor of Arts in Sociology, Shanghai Scholarship, Cumulative GPA:3.6

WORK EXPERIENCE

FancyDigital, Shanghai — Sep 2020 ~ Aug 2021
Data Analyst(Full-Time)

- Selected metrics and collaborated with a team to built real-time BI dashboards; Manipulated Clickhouse databases and visualized health of a SaaS system resulting in a 5% growth of customers
- Mentored new co-workers to perform user churn analysis based on users record data and provided user service suggestions to operators resulting in a decrease of user churn rate to 10%
- Cooperated with product managers to optimize product design and user experience using error logs data leading to a reduction of bug rate by 7%; Upgraded and tested data product in a SaaS system during developing schedule
- Designed ETL shell scripts to aggregate original detailed data into intermediate tables to optimize query performance by 85% and successfully automate daily reports; Reduced human-power input and time cost

QL2 Software, Baltimore, MD — Aug 2022 ~ Present
Data Analyst (Capstone)

- Extracted from MySQL databases and analyzed vehicle users data; Utilized Tableau to visualize features of different class of users; explored behavior and consumption habits of vehicle drivers
- Applied cluster analysis using Python to build vehicle user profile and provide marketing advice for dealers in automotive industry; Compiled report to present and interpreted statistical results

PROJECT EXPERIENCE

Click-Through Rate Prediction — Jul 2022

- Drill-down every feature to clarify importance using pandas and matplotlib; Performed data visualization and grouped Statistics; Selected features based on results of EDA to reduce complexity of training models
- Calculated click-through rate on high-dimensional features to separate data into bins, make every thousands-class feature into 5-class categorical factor, avoided sparse data problems and reduced program runtime by 15%
- Applied XGBoost, LightGBM and Sklearn in Python to build Predictive Models; found 11 variables associated with the expected results from a large dataset with 25 different variables

AWARDS AND HONORS

Kaggle M5 Forecasting Competition 103rd/5558 Top2% Silver Medal — Mar 2020 ~ Jun 2020
Estimate unit sales of Walmart retail goods

- Created features on tabular data using rolling-window method and trained a LightGBM algorithm to forecast daily sales for next 28 days; Built time series split to optimize hyper-parameters and acquired 0.58 of RMSE
- Stacked and averaged single LightGBM model predictions to improve the generalization ability of forecasts promoted 10% of ranking on leaderboard; Successfully avoided severe shake-down in privte leaderboard

TECHNICAL SKILLS

- Python (numpy, pandas, scikit-learn, tensorflow/keras), R (data.table, dplyr, shiny, tidyr)
- SQL(MySQL, SQL Server, TiDB, Clickhouse), Tableau, Linux/Shell, SPSS, Stata, Git, Web Crawler, Text Mining