QIUSHI YAN

qiushi.yan@vanderbilt.edu ⋄ 629 239-9151

Website ⋄ Github ⋄ Linkedin

EDUCATION

Vanderbilt University - Data Science Institute   **Nashville, TN**
Master of Science in Data Science   Expected May 2023

Communication University of China   **Beijing, China**
Bachelor of Arts in Data Journalism   September 2017 - June 2021

EXPERIENCE

RStudio   **Remote**
Software Engineer   Intern, May 2022 - August May 2022

- Developed the R package agua which provides tidymodels integration with the H2O machine learning platform.
- Contributed bug fixes and features to other tidymodels packages: parsnip, tune, usemodels, bundle.

51Talk   **Beijing, China**
Analytics Engineer   Intern, September 2020 - June 2021

- Identified key factors in teaching performance and ROI with regression modeling and hypothesis testing.
- Developed dashboards for monitoring teacher performance using metrics above with Django and Vue.

DiDi   **Beijing, China**
Data Mining Engineer   Intern, March 2019 - August 2019

- Created data cleaning pipeline and Shiny dashboards for AB testing and user research.

R PACKAGES

agua: tidymodels integration with H2O   package site

- Developed an R package that integrates the tidymodels ecosystem with the H2O machine learning platform, implemented functions for model fitting, tuning, parallel processing and reporting.

nyclodging: a analytical package for Airbnb listings   shiny app, package site

- Developed an R package and Shiny app to analyze Airbnb rentals in NYC, featuring spatial and text analysis of housing description. Used tidymodels and sf.

xkcd: xkcd-styled interactive charting library   package site

- Developed an interactive widget to create xkcd-styled plots. Used TypeScript and Esbuild for frontend tooling.

OTHER PROJECTS

Anime Recommendation System - Python

- Implemented a recommendation system with scraping and batch processing pipelines, rank and recall services, message queues and a custom Python package with service clients and common utilities.

Overwatch League Statistics API - Python, R   swagger doc, repo

- Created data ingestion and cleaning pipelines for analyzing player performance in the Overwatch League. Developed a restful API for querying transformed data. Used data.table and FastAPI.

ADDITIONAL

Languages	R. Python, TypeScript, Elixir
Frameworks	tidyverse, shiny, pyspark, FastAPI, tidymodels, data.table, React, Svelte
Certification	AWS Certified Developer