

QIUSHI YAN

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

[Website](#) ◇ [Github](#) ◇ [Linkedin](#)

EDUCATION

Vanderbilt University - Data Science Institute
Master of Science in Data Science

Nashville, TN
Expected May 2023

Communication University of China
Bachelor of Arts in Data Journalism

Beijing, China
September 2017 - June 2021

EXPERIENCE

RStudio - Software Engineer Intern

May 2022 - August 2022

- Developed the R package [agua](#) for integrating the H2O.ai machine learning platform with tidymodels.
- Contributed bug fixes and features to other R packages in tidymodels: parsnip, tune, usemodels, bundle.

51Talk - Analytics Engineer Intern

September 2020 - June 2021

- Modelled key variables in teaching performance and ROI with regression and hypothesis testing.
- Developed a dashboard for monitoring key metrics identified above with Django and Vue.

DiDi - Data Science Intern

March 2019 - August 2019

- Created data cleaning pipelines with PySpark and Shiny dashboards for visualizing AB testing results.

DATA SCIENCE PROJECTS

agua - Tidymodels Integration with H2O

[package site](#)

- Developed an R package providing tidymodels interface to the H2O sever, implemented a set of consistent APIs for model fitting, tuning, parallel processing and interpretation.

nyclodging - A Shiny app for Exploring Airbnb Listings

[shiny app](#), [package site](#)

- Developed an R package and Shiny app for analyzing Airbnb rentals in NYC. Deployed a multinomial classifier with textual features, enabling price predictions with arbitrary geolocation and housing descriptions.

Anime Recommendation - A recommendation system for anime series

- Implemented a complete ML product including scheduled ETL jobs with Airflow and PySpark, training pipelines using DeepWalk and Node2Vec, microservice GraphQL APIs and a MLOps dashboard.

SOFTWARE PROJECTS

qlang - A toy programming language with a combination of R and Python

[repo](#)

- A interpreter written in Go with a mix of features in R, Python and JavaScript. Provided native support for vectorization, easier syntax for creating arrays and hashmaps, and first-class functions.

POTG - A web application for sharing game clips

[website](#), [repo](#)

- An Angular + Firebase app for managing CRUD operations on video files. Used webassembly for video processing. Implemented a simple graph-based recommendation engine.

For more projects see [projects.qiushiyan.dev](#).

ADDITIONAL

Frameworks Tidyverse, Shiny, FastAPI, Airflow, Vue, React
Certification [AWS Certified Developer](#)